УДК 575.113/.118
ББК 28.04-28.070
Л91

Серия основана в 2006 г.

Льюин Б.

Л91 Гены / Б. Льюин ; пер. 9-го англ. изд. — М. : БИНОМ. Лаборатория знаний, 2012. — 896 с. : цв. ил. — (Лучший зарубежный учебник).

ISBN 978-5-94774-793-5

Известный автор Бенджамин Льюин подготовил новое издание книги, ставшей классикой для молекулярных биологов всего мира. Его обстоятельная работа всегда соответствует последним достижениям в области молекулярной биологии и молекулярной генетики, включая структуру генов, последовательности, организацию и экспрессию. Девятое издание дополнено новыми разделами (эпигенетические эффекты), хорошо иллюстрировано и структурировано, что помогает студентам лучше ориентироваться в отдельных темах.

Для студентов, специализирующихся в области молекулярной генетики, молекулярной биологии, генной инженерии, а также для аспирантов, преподавателей, научных сотрудников.

УДК 575.113/.118
ББК 28.04-28.070

Учебное издание

Серия: «Лучший зарубежный учебник»

Льюин Бенджамин

ГЕНЫ

Ведущий редактор канд. биол. наук *В. В. Гейдебрехт*
Редактор *А. В. Парамонова*
Художник *Н. А. Новак*
Технический редактор *Е. В. Денюкова*
Корректор *Е. Н. Клитина*
Компьютерная верстка: *Т. Э. Внукова*

Подписано в печать 28.09.11. Формат 60×90/8.
Усл. печ. л. 112. Тираж 1500 экз. Заказ

Издательство «БИНОМ. Лаборатория знаний»
125167, Москва, проезд Аэропорта, д. 3
Телефон: (499) 157-5272, e-mail: binom@Lbz.ru, http://www.Lbz.ru

Отпечатано в Венгрии при участии фирмы «ИНТЕРПРЕСС»
interpress@interpress.eu

ORIGINAL ENGLISH LANGUAGE EDITION PUBLISHED BY
Jones and Bartlett Publishers, Inc.
40 Tall Pine Drive
Sudbury, MA 01776

ISBN 978-5-94774-793-5

Предисловие редактора перевода

Первое издание книги Бенджамина Льюина «Гены» увидело свет в 1983 г. А в 1987 г. книга стала доступна на русском языке. С тех пор англоязычная версия «Генов» пережила восемь переизданий, перевод последнего из которых — девятого — решило опубликовать издательство «БИНОМ. Лаборатория знаний».

В современном информационном мире, когда у каждого в кармане находится практически весь запас накопленных человечеством знаний (с хорошо работающими системами быстрого поиска любого определения, значения или формулы), главной ценностью таких изданий, как «Гены» становится эффективное структурирование огромного объема научных данных и их подача в удобной для читателя форме и последовательности. Структура книги с развернутым описанием содержания каждой главы позволяет быстро сориентироваться в тексте.

Не секрет, что язык современного отечественного молекулярного биолога практически полностью состоит из англоязычных терминов, для большинства из которых даже не придуманы русскоязычные аналоги.

При переводе и редактировании насыщенного научными терминами англоязычного текста всегда сложно определить ту грань, до которой их еще нужно заменять отечественными аналогами, а после которой — целесообразно давать лишь кириллическую транскрипцию. Мы постарались угадать этот рубеж, хотя среди читателей разных поколений, безусловно, найдутся критики как недорусифицированности, так и консервативности перевода. Употребление в ряде случаев терминов, отличающихся от использованных в первом отечественном издании «Генов» (под редакцией Г. П. Георгиева), вызвано постоянным изменением языка практической науки, бороться с которым на наш взгляд нецелесообразно.

Книгу переводили И. А. Кофиади (гл. 1–7), Н. Ю. Усман (гл. 8, 14–22, 25, 27–31, словарь терминов), М. А. Турчанинова (гл. 9–13, 26) и А. М. Савилова (гл. 23, 24).

Ребриков Д. В.

Предисловие

Наука — удивительно увлекательна и одновременно рискованна. Каждое новое издание этой книги дополняется интереснейшими открытиями, и то, которое вы держите в руках, включает обновленные данные исследований в области молекулярной биологии. Структура данного издания была переработана так, чтобы соответствовать структуре изложения материала в книге «Незаменимые гены», облегчая использование обеих книг в связке. Поскольку я старался избежать существенного увеличения объема и без того немаленькой книги, в данном издании за счет сокращения глав, повествующих о следствиях работы генов и цитологических аспектах строения клетки, материал сфокусирован на собственно генах и их экспрессии. Наибольшие изменения произошли в первой части книги, описывающей строение геномов, что явилось отражением бурного развития полногеномного секвенирования. Важность РНК как генетического регулятора стала еще более очевидной, распространившись на все уровни экспрессии генов как для про-, так и для эукариот. Являясь в некотором роде «недостающим звеном эволюции», РНК проливает свет на то, как современный аппарат экспрессии генов мог возникнуть из древнейших РНК-опосредованных живых систем.

Моя стратегия при написании данной книги состояла в цитировании оригинальных публикаций и обзоров, доступных читателю. Предпочтение отдавалось статьям, выходящим в свободный доступ через полгода после публикации, а там, где это не удавалось, отбирал наиболее доступные публикации.

Я выражаю благодарность следующим моим коллегам, выступившим в качестве корректоров и консультантов данного издания:

Эллиотт Гольдшейн	Университет штата Аризона, Тэмпе;
Джоселин Крэбс	Университет штата Аляска; Анкоридж;
Кэтлин Мэтьюз	Университет Райса, Хьюстон.

Бенджамин Льюин
Январь 2007

Структура книги

Новая организация издания *ГЕНЫ IX* позволяет преподавателям и студентам детальнее сфокусироваться на генах и их экспрессии за счет более широко рассмотрения ключевых аспектов. Число глав и последовательность разделов остались прежними, однако некоторые главы были подразделены на две или более главы. Перечислим отличия данного издания от предыдущего:

Глава 1 издания *ГЕНЫ VIII* («Гены — это ДНК») в текущем издании разбита на две главы. Информация по структуре ДНК, репликации и мутационным процессам осталась в главе 1, в то время как рассмотрение функции генов как элемента наследственности вынесено в новую («Гены кодируют белки»).

Глава 3 издания *ГЕНЫ VIII* («Содержимое генома») разбита в данном издании на две главы.

Глава 4 данного издания («Структура генома») включает информацию о последовательностях ДНК, картировании геномов и ДНК органелл.

Глава 5 («Последовательности генома и число генов») теперь включает данные о размерах геномов и экспрессии генов для многих организмов, а также новые данные о генах Y-хромосомы.

Новая глава 12 («Оперон») включает в себя главу 10 издания *ГЕНЫ VIII*, а также информацию по регуляции транскрипции и трансляции из главы 11 («Регуляторные циклы»). Материалы по регуляторным РНК перенесены в главу 13.

Материалы главы 13 издания *ГЕНЫ VIII* («Репликон») расширены до трех глав. Глава 15 («Репликон») текущего издания описывает структуру и функции репликона и ориджинов репликации. Глава 16 («Внехромосомные репликоны») содержит информацию о терминальных белках, репликации по типу катящегося кольца, плазмидам и Т-ДНК. Информация о взаимосвязи бактериальной репликации с клеточным циклом можно найти в главе 17.

Глава 15 издания *ГЕНЫ VIII* («Рекомбинация и репарация») теперь разбита на две главы. Глава 19 описывает гомологичную и сайт-специфичную рекомбинацию, а глава 20 — системы репарации, включая работу систем эксцизионной репарации в клетках млекопитающих.

Глава 23 издания *ГЕНЫ VIII* («Контроль структуры хроматина») теперь идет под номером 30, описывая взаимосвязь структуры хроматина и экспрессии генов.

Глава 31 («Эпигенетические эффекты наследуются») описывает особенности механизмов эпигенетического наследования и их последствия.

Иллюстрации и дизайн

Издание *ГЕНЫ IX* выглядит по-новому. Дизайн и иллюстрации были обновлены и дополнены, чтобы облегчить понимание материала студентами. Кроме того, стиль и дизайн издания *ГЕНЫ IX* ориентирован на новую версию книги «Клетки» (Б. Льюина), позволяя студентам и преподавателям использовать обе книги в качестве единого модуля.

Приложения

Для студентов

Специально созданный в рамках выхода девятого издания книги Интернет-сайт (http://biology.jbpub.com/book/genes) содержит множество ресурсов, помогающих лучше разобраться в такой области знаний, как молекулярная биология.

Лабораторные исследования в молекулярной биологии (Вильямс, Слатко и МакКерри) содержат широко используемые в исследовательских лабораториях молекулярно-биологические протоколы. Представленные эксперименты спланированы так, чтобы дать студентам ощущение настоящего научного проекта, выполняемого на современной исследовательской базе.

Для преподавателей

Совместимый с Windows и Macintosh «Набор преподавателя» (CD-ROM) (в английском варианте издания. — *Прим. ред.*) дает возможность лектору применять в процессе обучения следующие элементы:

- Банк тестов доступен в виде выверенных текстовых файлов.
- Лекционные слайды, выполненные для программы PowerPoint, содержат иллюстрации к каждой главе издания *ГЕНЫ IX*. Лекторы, использующие Microsoft PowerPoint, могут самостоятельно оптимизировать содержание, оформление и последовательность слайдов презентации.
- Коллекция изображений содержит ключевые иллюстрации и таблицы, представленные в тексте книги (на которые издательство Jones and Bartlett Publishers сохраняет права на публикацию или цифровое воспроизведение). Библиотека изображений позволяет лектору демонстрировать иллюстрации, приведенные в книге, в аудитории, добавляя их в презентацию.

Сокращения

А	Аденин или аденозин
АДФ	Аденозиндифосфат
АМФ	Аденозинмонофосфат
цАМФ	Циклический АМФ
АТФ	Аденозинтрифосфат
АТФаза	Аденозинтрифосфатаза
пн	Пары нуклеотидов
Ц	Цитидин или цитозин
кДНК	Комплементарная ДНК
ЦДФ	Цитидиндифосфат
ЦМФ	Цитидинмонофосфат
ЦТФ	Цитидинтрифосфат
ДНК	Дезоксирибонуклеиновая кислота
ДНКаза	Дезоксирибонуклеаза
Г	Гуанин или гуанозин
ГДФ	Гуанозиндифосфат
GlcNAc	N-ацетил-D-глюкозамин
ГМФ	Гуанозинмонофосфат
ГТФ	Гуанозинтрифосфат
ΔG	Изменение свободной энергии
тпн	Тысячи пар нуклеотидов
Мпн	Миллионы пар нуклеотидов
иРНК	Информационная РНК
MW	Молекулярная масса
Pi	Неорганический фосфат
Ppi	Неорганический пирофосфат
РНК	Рибонуклеиновая кислота
РНКаза	Рибонуклеаза
рРНК	Рибосомальная РНК
тРНК	Транспортная РНК
Т	Тимин или тимидин
У	Урацил
УДФ	Уридиндифосфат
УМФ	Уридинмонофосфат
УТФ	Уридинтрифосфат

Обозначение	**Кратность**
Мега (М)	10^{6}
Кило (к)	10^{3}
Деци (д)	10^{-1}
Санти (с)	10^{-2}
Милли (м)	10^{-3}
Микро (мк)	10^{-6}
Нано (н)	10^{-9}
Пико (п)	10^{-12}

Единицы	
Å	Ангстрем
Да	Дальтон
г	Грамм
ч	Час
М	Молярная концентрация
м	Метр
мин	Минута
N	Ньютон
S	Единицы Сведберга
с	Секунда
V	Вольт

Аминокислоты		
A	Ala	Аланин
C	Cys	Цистеин
D	Asp	Аспарагиновая кислота
E	Glu	Глутаминовая кислота
F	Phe	Фенилаланин
G	Gly	Глицин
H	His	Гистидин
I	Ile	Изолейцин
K	Lys	Лизин
L	Leu	Лейцин
M	Met	Метионин
N	Asn	Аспарагин
P	Pro	Пролин
Q	Gln	Глутамин
R	Arg	Аргинин
S	Ser	Серин
T	Thr	Треонин
V	Val	Валин
W	Trp	Триптофан
Y	Tyr	Тирозин

Гены — это ДНК

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

1.10 Нуклеиновые кислоты гибридизуются в соответствии с принципом комплементарности

- Нагревание приводит к разделению цепей ДНК-дуплекса.
- Температура плавления — это температура разделения цепей дуплекса.
- Комплементарные одноцепочечные молекулы ДНК при понижении температуры могут ренатурировать (с образованием дуплекса).
- Процесс денатурации (плавления) и ренатурации (гибридизации) может происходить как между однотипными молекулами, так и между молекулами разного типа. При этом могут формироваться дуплексы ДНК-ДНК, РНК-ДНК и РНК-РНК.
- Способность двух одноцепочечных нуклеиновых кислот гибридизоваться зависит от степени их комплементарности.

1.11 Мутации изменяют последовательность ДНК

- Все мутации — это изменения в последовательности ДНК.
- Мутации могут возникать спонтанно или могут быть спровоцированы мутагенами.

1.12 Мутации могут затрагивать отдельные пары оснований или более длинные последовательности

- Точечные мутации изменяют отдельные пары оснований.
- Точечные мутации могут быть результатом химического превращения одного основания в другое или могут возникать в результате ошибок репликации.
- Замену пары G-C на A-T (или наоборот) называют транзицией.
- Замену пурина на пиримидин (например, A-T на T-A) называют трансверсией.
- Инсерции (вставки) — самый распространенный тип мутаций — происходят в результате перемещения мобильных элементов.

1.13 Эффекты мутаций могут быть обратимы

- Прямые мутации инактивируют ген, а обратные мутации (реверсии) способны компенсировать эти эффекты.
- Эффект от инсерции может быть обращен делецией вставленного участка, тогда как делеция практически необратима.
- Когда мутация в одном гене сводит на нет эффект от мутации в другом гене, говорят об эффекте подавления (супрессии).

1.14 Мутации наиболее часты в «горячих точках» генома

- На большей части генома вероятность возникновения мутаций примерно одинакова, однако в «горячих точках» эта вероятность может возрастать на порядки.

1.15 Большинство «горячих точек» мутагенеза появляется в результате модификации оснований

- Спонтанное дезаминирование, в результате которого происходит превращение 5-метилцитозина в тимин, — одна из наиболее частых причин возникновения «горячих точек».

1.16 Некоторые элементы наследственности чрезвычайно малы

- Некоторые очень маленькие элементы наследственности не кодируют собственных белков, однако, представляя собой РНК или белок, обладают наследственным потенциалом.

1.17 Резюме

1.1 Введение

Способность передавать признаки по наследству является естественным свойством любого живого организма. Наследственная *информация*, необходимая для создания нового организма, содержится в **нуклеиновых кислотах**, совокупность которых называют **геномом**. Мы используем термин «информация», поскольку геном как таковой не принимает непосредственного участия в построении организма; точнее будет сказать, что он представляет собой некую последовательность субъединиц (оснований), формирующих нуклеиновые кислоты, которые определяют наследственные признаки. В результате ряда сложных взаимодействий эта последовательность принимает участие в формировании всего разнообразия белков организма, синтезируя их в нужное время и в нужном месте. Белки, в свою очередь, либо формируют собственно структуру клетки и организма (структурные белки), либо принимают участие в метаболических реакциях, необходимых для поддержания жизни организма.

Геном любого живого организма содержит исчерпывающую информацию о признаках, передающихся по наследству. Физически геном состоит из набора нуклеиновых кислот, а функционально он состоит из **генов**. Каждый ген представлен последовательностью нуклеотидов некоторого участка нуклеиновой кислоты и кодирует информацию об одном белке. Отдельная молекула нуклеиновой кислоты из общего разнообразия, составляющего геном, может содержать значительное количество генов. Размеры геномов живых организмов варьируют в широком диапазоне. Так, например, геном бактерии *Micoplasma* состоит всего из 500 генов, а геном человека из более чем 25 000.

В этой части книги мы попытаемся проанализировать свойства гена, основываясь на базовых представлениях о его молекулярной организации. На РИС. 1.1 приведены основные исторические этапы развития концепции гена, начиная с первых теоретических воззрений и заканчивая современными данными о структуре генома.

Каждый отдельно взятый геном представлен полным набором хромосом данного организма. Таким образом, геном состоит из разных молекул ДНК (по одной на каждую хромосому), каждая из которых, в свою очередь, содержит множество генов. Если попытаться дать геному буквальное определение, то это будет совокупность последовательностей всех молекул ДНК, из которых состоят хромосомы.



РИС. 1.1 Краткая история генетики

Первое представление о гене как о функциональной единице появилось после доказательства того, что каждый отдельно взятый ген ответственен за производство определенного белка. Различия в химической структуре ДНК, из которой состоит ген, и химической структуре белков послужили развитию концепции генетического кода. Это, в свою очередь, привело к открытию сложного аппарата, осуществляющего перевод последовательности ДНК в последовательность аминокислот полипептидной цепи белка.

С пониманием процесса экспрессии гена, представления о его структуре и функциях стали значительно более полными. На **РИС. 1.2** продемонстрирован основной предмет обсуждения данной книги.

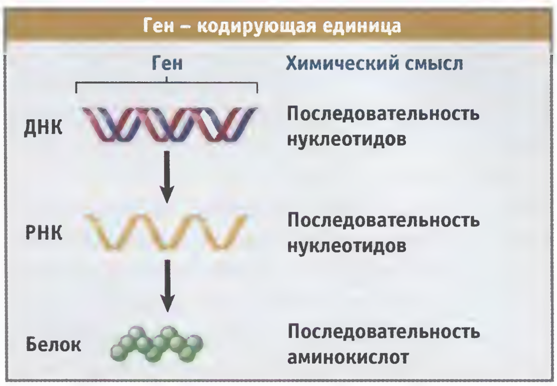


РИС. 1.2 Гены кодируют РНК, которая, в свою очередь, может кодировать белок

Ген — это фрагмент последовательности ДНК, в котором закодирована информация о структуре другой нуклеиновой кислоты — РНК. ДНК — двухцепочечная нуклеиновая кислота, тогда как РНК, как правило, представлена одной цепью. Последовательность РНК определяется последовательностью ДНК (фактически, РНК идентична одной из цепей ДНК). В большинстве случаев (но не во всех) РНК принимает непосредственное участие в производстве белка. *Таким образом, ген — это последовательность ДНК, которая кодирует РНК; РНК, в свою очередь, несет информацию о белке (если, конечно, она является продуктом гена, кодирующего белок).*

От иллюстрации того, что ген представляет собой фрагмент ДНК, а хромосомы — это длинные цепи ДНК, содержащие множество генов, мы перейдем к рассмотрению общей организации генома, но уже с позиции последовательности ДНК. В третьей главе («Прерывистый ген») мы более детально рассмотрим строение гена и его участие в синтезе белка. В четвертой главе («Состав генома») мы рассмотрим все разнообразие генов, и, наконец, в шестой главе («Кластеры и повторы») обсудим другие компоненты генома и механизмы поддержания его целостности.

1.2 ДНК — генетический материал бактерий

Ключевые понятия:

- Первым доказательством того, что генетический материал бактерий представляет собой ДНК, стало открытие бактериальной трансформации. Было показано, что перенос ДНК от одного штамма к другому приводит к передаче генетических признаков.

Открытие процесса бактериальной трансформации в 1928 г. послужило развитию идеи о том, что нуклеиновые кислоты служат генетическим материалом. Бактерия *Pneumococcus* вызывает у мышей пневмонию. Вирулентность этой бактерии обусловлена *капсульным полисахаридом*. Этот компонент клеточной стенки позволяет бактерии избегать воздействия защитной системы организма. Разные типы пневмококков (I, II и III) содержат различные капсульные полисахариды. Они образуют гладкие колонии, обозначаемые буквой S (от *англ.* smooth — «гладкий»).

Каждый пневмококк S-типа может производить дочерние бактерии, в которых механизм производства полисахарида оболочки нарушен. Эти бактерии обозначают буквой R (от *англ.* rough — «грубый») и имеют шероховатую поверхность (состоящую из материала, расположенного под полисахаридом оболочки). Эти пневмококки **авирулентны**. Они не способны вызвать заболевание, поскольку недостаток полисахарида делает их уязвимыми.



РИС. 1.3 Ни термически инактивированные бактерии S-типа, ни живые R-типа не могут убить мышь, однако их одновременная инъекция убивает мышей так же эффективно, как живые бактерии S-типа



РИС. 1.4 ДНК бактерий S-типа может трансформировать бактерии R-типа в S-тип

Если подвергнуть бактерии S-типа температурному воздействию, они гибнут и теряют способность причинять мышам вред. Однако погибшие S-бактерии вместе с непатогенными живыми R-бактериями демонстрируют совсем иные свойства, нежели каждая из них по отдельности. На **РИС. 1.3** показано, что при инъекции смеси таких бактерий животное вскоре погибает от пневмонии, а после смерти из мышей можно выделить патогенные бактерии S-типа.

В одном из экспериментов исследователи подвергли пневмококков III типа температурному воздействию, а живых R-бактерий получили из пневмококков II типа. Оказалось, что патогенные бактерии, выделенные из мертвых мышей, подвергшихся воздействию такой смешанной инфекции, оказались покрыты оболочкой, содержащей полисахариды, характерные для III типа. Таким образом, инактивированные патогенные пневмококки III типа *трансформировали* (изменили) живых бактерий II типа, наделив их способностью производить полисахариды III типа.

На **РИС. 1.4** показан процесс выделения компонента инактивированных бактерий, ответственного за трансформацию. Этот компонент был назван **трансформирующим фактором**. Он был очищен благодаря созданию бесклеточной системы, в которой фактор, выделенный из инактивированной бактерии S-типа можно было добавлять к живым бактериям R-типа до введения в мышь. В результате очистки трансформирующего фактора в 1944 году было показано, что это **дезоксирибонуклеиновая кислота (ДНК)**.

1.3 ДНК — генетический материал вирусов

Ключевые понятия:

- Доказательство того, что ДНК — это генетический материал некоторых вирусов, было получено в результате изучения процесса инфицирования бактерий фагами. Если пометить разными радиоактивными метками ДНК и белки бактериофага, можно обнаружить, что дочерним фагам, продуцируемым инфицированной бактерией, передается только ДНК.

После эксперимента с бактериями, показавшего, что их генетический материал представлен ДНК, исследователи продемонстрировали, что ДНК служит источником генетической информации и в совершенно иной системе. Фаг T2 — это вирус, способный заражать бактерию *Escherichia coli*. При добавлении фага (вируса) в культуру *E. coli* его частицы адсорбируются на внешней поверхности бактерий, часть вирусного материала проникает внутрь клетки, после чего, примерно через 20 мин, бактерия погибает (лизируется), высвобождая огромное количество дочерних фагов.

На **РИС. 1.5** показаны результаты эксперимента по заражению *E. coli* фагом T2, проведенного в 1952 г. В эксперименте разные компоненты фага были помечены различными радиоактивными метками. ДНК была помечена изотопом ^{32}P, а белки — изотопом ^{35}S. Смесь бактерий при помощи центрифугирования разделили на две фракции. Одна фракция содержала пустые фаговые оболочки с поверхности *E. coli*, а другая содержала инфицированных бактерий.

Метка ^{32}P (метка ДНК) большей частью оказалась во фракции с инфицированными бактериями. Дочерние фаги содержали около 30% от исходного количества метки ^{32}P и менее 1% от исходного количества меченого белка. Таким образом, эксперимент продемонстрировал, что в бактерию попадает только ДНК материнского фага, которая затем становится частью дочерних фагов. Эта модель продемонстрировала передачу генетического материала фагов по наследству.

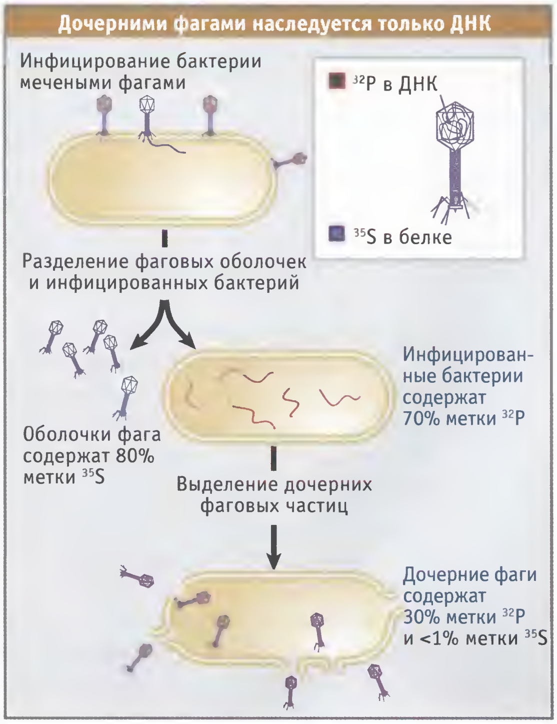


РИС. 1.5 Генетический материал фага Т2 представлен ДНК

Размножение фага происходит за счет ресурсов клетки хозяина. Фаг обладает генетическим материалом, аналогичным генетическому материалу клетки: его особенности точно воспроизводятся и подчиняются тем же универсальным правилам. Эксперимент с фагом Т2 подтверждает основной постулат: ДНК — это генетический материал, вне зависимости от того, является она геномом клетки или вируса.

1.4 ДНК — генетический материал клеток животных

Ключевые понятия:

- ДНК можно использовать для придания новых признаков как отдельным клеткам животных, так и целым организмам.
- Генетический материал некоторых вирусов представлен РНК.

Если добавить ДНК к культуре эукариотических клеток, нуклеиновые кислоты могут проникать внутрь клеток и в некоторых из них могут вызывать синтез новых белков. Если при этом использовать фрагмент ДНК, несущий конкретный ген, его проникновение в клетку приводит к синтезу закодированного белка. На **РИС. 1.6** изображена одна из таких систем.

Несмотря на то что по историческим причинам процедура введения ДНК в клетки эукариот была названа **трансфекцией**, ее механизм в целом аналогичен бактериальной трансформации. ДНК, проникая в клетку-реципиент, становится частью ее генетического материала, и привнесенная информация передается по наследству так же, как это происходит с любой другой частью ДНК. Экспрессия нового участка ДНК придает клеткам новые свойства (в случае, изображенном на рис. 1.6, это способность синтезировать тимидинкиназу). Сначала такие эксперименты удавались только с индивидуальными клетками, выращенными в культуре. Однако впоследствии удалось ввести ДНК в яйцеклетки мышей и добиться того, чтобы введенная ДНК стала частью их генетического материала.

На основании таких экспериментов было доказано, что ДНК — генетический материал эукариот и что *при межвидовом переносе нуклеиновые кислоты могут сохранять способность выполнять свои функции.*

Как мы уже выяснили, ДНК — это генетический материал всех известных клеточных организмов и многих вирусов. Однако у некоторых вирусов генетический материал представлен *рибонуклеиновой кислотой* (РНК). Таким образом, основной принцип природы генетического материала состоит в том, что это всегда нуклеиновая кислота; а точнее — это всегда ДНК, за исключением РНК-содержащих вирусов.

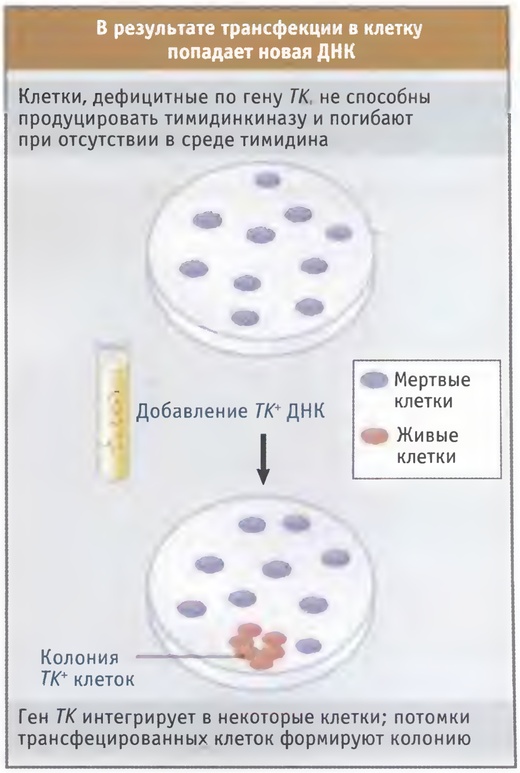


РИС. 1.6 При добавлении ДНК в культуру эукариотических клеток происходит трансфекция, в результате которой клетка может приобрести новый фенотип

1.5 Полинуклеотидные цепи состоят из азотистых оснований, связанных с сахарофосфатным остовом

Ключевые понятия:

- Нуклеозиды состоят из пуриновых или пиримидиновых оснований, связанных в первой позиции с пятиуглеродным сахаром (пентозой).
- Чтобы различать позиции атомов углерода, входящих в состав азотистого основания и пентозного кольца, атомы сахара обозначают штрихом (′).
- ДНК и РНК отличаются химической группой в положении 2′ сахарного кольца: в ДНК — это водород (что дает дезоксирибозу), а в РНК — гидроксил (давая рибозу).
- Нуклеотид состоит из нуклеозида, связанного с фосфатной группой в положении 5′- или 3′- (дезокси)рибозы.
- Каждая последующая субъединица полинуклеотидной цепи соединяется своей 5′-позицией через фосфатную группу с 3′-позицией предыдущего сахара.
- Один из концов цепи называют 5′-конец, а противоположный — 3′-конец.
- ДНК содержит четыре основания — аденин, гуанин, цитозин и тимин; в РНК вместо тимина — урацил.

Основной составной частью нуклеиновой кислоты является нуклеотид. Нуклеотид состоит из трех компонентов:

- азотистое основание;
- сахар (пентоза) и
- фосфат.

Азотистое основание содержит пуриновое или пиримидиновое кольцо и связано гликозидной связью (через N_9 у пуринов и N_1 у пиримидинов) с 1′-положением пятиуглеродного сахара. Чтобы избежать неясностей с нумерацией атомов гетероциклических колец и атомов сахара, позиции пентозного кольца отмечают штрихом (′).

Тип сахара, включенного в состав нуклеотида, определяет название нуклеиновой кислоты. В нуклеотидах, составляющих ДНК, содержится 2′-дезоксирибоза, а в РНК — рибоза. Разница между рибозой и 2′-дезоксирибозой состоит в том, что в положении 2′ у рибозы находится OH-группа. Сахар может быть связан с фосфатной группой в 5′- или 3′-положении.

Нуклеиновая кислота состоит из длинной цепи нуклеотидов. На РИС. 1.7 показано, что остовом полинуклеотидной цепи служит чередующаяся последовательность пятиуглеродных сахаров и фосфатов. Этот остов формируется путем соединения 5′-позиции одного сахара и 3′-позиции другого сахара через фосфатную группу. Таким образом, сахарофосфатный остов состоит из 5′–3′ фосфодиэфирных связей. Азотистые же основания обращены в сторону от остова.

Каждая нуклеиновая кислота состоит из четырех типов оснований: 2 пуриновых и 2 пиримидиновых. Пуриновые основания — аденин и гуанин присутствуют и в ДНК и в РНК. Что же касается пиримидинов, то в состав ДНК входят цитозин и тимин, а в РНК вместо тимина содержится урацил. Единственное различие между урацилом и тимином в том, что у урацила в позиции C_5 находится метильная группа. Основания обычно обозначают по первым буквам их названий: в ДНК — A, G, C и T; в РНК — A, G, C и U. У крайнего нуклеотида с одной стороны цепи находится свободная 5′-группа, а у крайнего с другой стороны — 3′-группа. Записывать последовательность нуклеиновых кислот принято в направлении от 5′- к 3′-концу, при этом располагая 5′-конец слева, а 3′-конец справа.

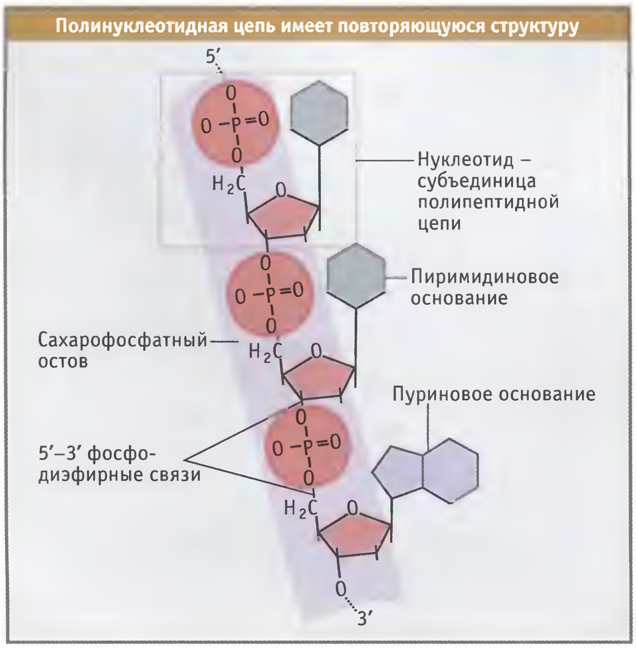


РИС. 1.7 Полинуклеотидная цепь состоит из серии 5′-3′ сахарофосфатных связей, формирующих остов, из которого «торчат» основания

1.6 Структура ДНК представлена двойной спиралью

Ключевые понятия:

- В-форма ДНК представлена двойной спиралью, состоящей из двух антипараллельных полинуклеотидных цепей.
- Азотистые основания каждой цепи состоят из плоских пуриновых или пиримидиновых колец обращенных друг к другу и соединенных посредством водородных связей в пары A-T или G-C.
- Диаметр двойной спирали — 20 Å. Полный оборот спирали включает 10 оснований и происходит на промежутке длиной в 34 Å.
- Двойная спираль формирует основную (большую) и второстепенную (малую) бороздку.

У различных видов живых существ основания, входящие в состав нуклеиновых кислот, представлены в разных соотношениях. Это наблюдение привело к развитию концепции, в соответствии с которой генетическая информация определяется последовательностью оснований. В 1950-х годах эта концепция нашла подтверж-

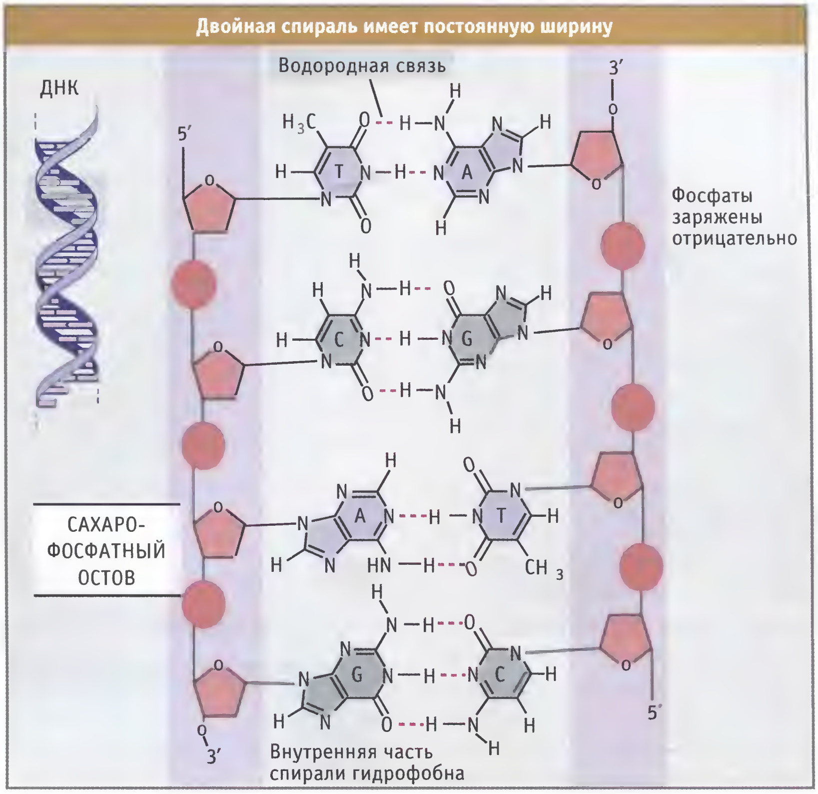


РИС. 1.8 Двойная спираль сохраняет постоянную ширину, поскольку в комплементарных парах A-T и G-C пурины всегда противостоят пиримидинам. На рисунке представлена последовательность TCAG

дение: разработанная модель структуры ДНК объясняла, как последовательность оснований ДНК может определять последовательность аминокислот в белке.

Предложенная в 1953 г. Уотсоном и Криком модель ДНК, состоящая из двух цепей, принимала в расчет результаты трех наблюдений:

- Данные по преломлению рентгеновских лучей показали, что ДНК имеет форму правильной спирали. Спираль совершает полный оборот каждые 34 Å (3,4 нм), а ее диаметр равен 20 Å (2 нм). Поскольку расстояние между соседними нуклеотидами составляет 3,4 Å, на один полный оборот спирали должно приходиться 10 нуклеотидов.
- Плотность молекулы ДНК позволяет предположить, что спираль состоит из двух полинуклеотидных цепей. Спираль на всем своем протяжении имеет постоянный диаметр. Это можно объяснить тем, что основания каждой цепи обращены друг к другу и расположены так, что пурин всегда лежит напротив пиримидина. Пары пурин-пурин слишком велики, для того чтобы обеспечить постоянство диаметра, а пары пиримидин-пиримидин — слишком малы.
- Вне зависимости от абсолютного размера каждого основания, доля G всегда равна доле основания C. То же справедливо для соотношения A и T. Таким образом, состав ДНК может быть описан долей ее оснований: G+C. Она варьирует в интервале от 26% до 74% у разных видов.

Уотсон и Крик предположили, что две полинуклеотидные цепи в двойной спирали соединяются водородными связями между их азотистыми основаниями. G может связываться только с C, а A только с T. Такое взаимодействие называют **спариванием оснований**, а спаренные таким образом основания принято называть **комплементарными**.

В соответствии с предложенной моделью две полинуклеотидные цепи противоположно направлены (**антипараллельны**) (РИС. 1.8). На протяжении всей спирали одна из цепей лежит в направлении от 5′-конца к 3′-концу, тогда как комплементарная ей цепь направлена в противоположную сторону.

Сахарофосфатный остов расположен с внешней стороны спирали и несет отрицательные заряды на фосфатных группах. В растворе *in vitro* отрицательные заряды нейтрализованы ионами металлов, в большинстве случаев Na^{+}. В клетке нейтрализующую функцию выполняют положительно заряженные белки. Эти белки играют важную роль при формировании структуры ДНК в клетке.

Основания направлены внутрь спирали. Это плоские структуры, расположенные попарно перпендикулярно оси спирали. Представьте себе двойную спираль ДНК в виде винтовой лестницы, у которой ступеньки — это пары оснований (РИС. 1.9). Поднимаясь вдоль спирали, основания укладываются одно за другим как стопка тарелок.

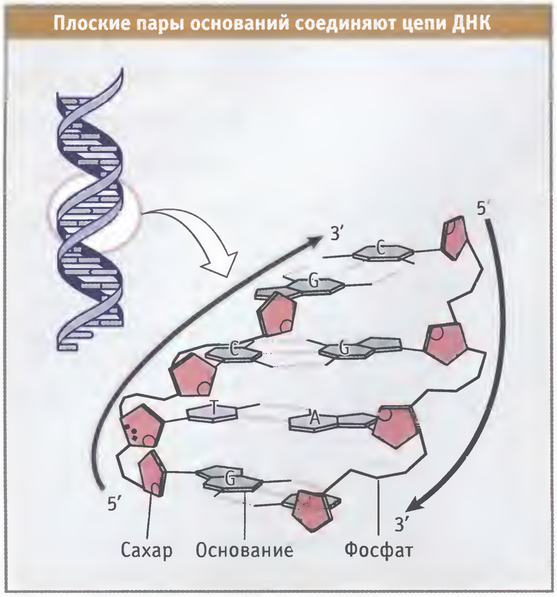


РИС. 1.9 Плоскость, в которой лежат пары оснований, расположена перпендикулярно сахарофосфатному остову

Каждая пара оснований повернута на 36° вокруг оси спирали относительно следующей пары. Таким образом, 10 пар оснований делают полный оборот в 360°. Две цепи, закручиваясь друг относительно друга, формируют **малую бороздку** (около 12 Å в поперечнике) и **большую бороздку** (около 22 Å в поперечнике) (РИС. 1.10). Двойная спираль закручена **вправо**, то есть

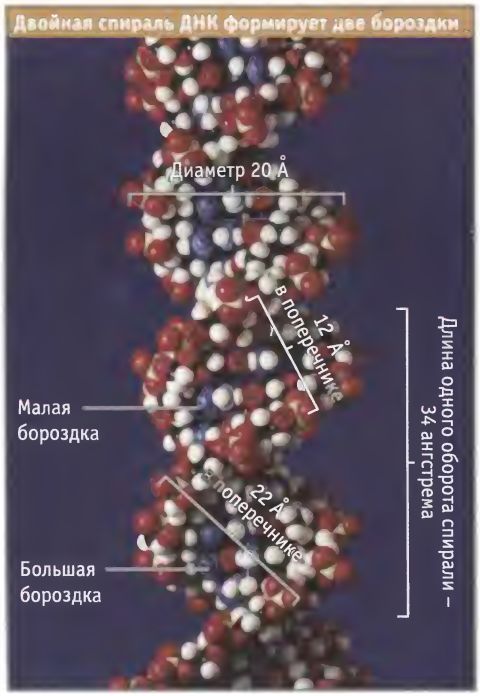


РИС. 1.10 Две цепи ДНК формируют двойную спираль. Фото © Photodisc

ее ход соответствует ходу часовой стрелки, если взглянуть вдоль оси спирали. Приведенные выше параметры соответствуют **В-форме** ДНК.

Важно понимать, что В-форма отражает некое усредненное представление о структуре ДНК. Ее структура может изменяться от участка к участку. ДНК может содержать большее число пар оснований на оборот спирали, и тогда говорят, что она **перекручена**; если же число оснований на виток спирали меньше 10, говорят, что она **недокручена**. Эти особенности могут быть результатом пространственных изменений в структуре двойной спирали или могут быть связаны с прикреплением белковых молекул к некоторым участкам ДНК.

1.7 Процесс репликации ДНК полуконсервативен

Ключевые понятия:

- Эксперимент Мезельсона–Сталя, основанный на мечении ДНК изотопами, доказал, что в реакции репликации консервативным элементом ДНК служит одна из полинуклеотидных цепей.
- Каждая цепь ДНК-дуплекса служит матрицей для синтеза дочерней цепи.
- Последовательность нуклеотидов, составляющих дочернюю цепь, комплементарна последовательности материнской цепи.

Важным свойством генетического материала является возможность его точного воспроизводства. Две полинуклеотидные цепи связаны лишь водородными связями, а значит, их можно легко разъединить, поскольку для этого не требуется разрушения прочных ковалентных связей. Специфичность спаривания оснований позволяет сделать предположение, что каждая из разделенных цепей **материнской** ДНК может служить **матрицей** для синтеза комплементарной **дочерней** цепи. На РИС. 1.11 показан принцип, в соответствии с которым каждая новая дочерняя цепь собирается на материнской цепи. Последовательность дочерней цепи определяется материнской цепью. Аденин в материнской цепи приводит к встраиванию тимина в дочернюю цепь, а гуанин определяет встраивание в дочернюю цепь цитозина и т. п.

В верхней части рис. 1.11 изображен материнский (нереплицированный) ДНК-дуплекс, изначально состоящий из двух материнских цепей. На нижней части рисунка изображены два дочерних дуплекса, появившиеся в результате присоединения комплементарных оснований. Каждый из дочерних дуплексов идентичен последовательности исходного материнского и состоит из одной материнской цепочки и одной вновь синтезированной. *Структура ДНК позволяет сохранять ее последовательность при репликации.*

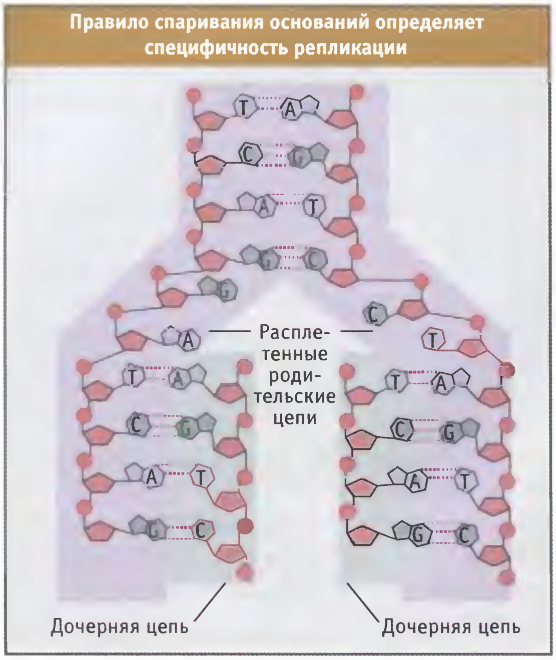


РИС. 1.11 Репликация ДНК осуществляется в соответствии с правилом спаривания оснований

Результат такого способа репликации изображен на **РИС. 1.12**. Материнский дуплекс реплицируется, давая два дочерних дуплекса, каждый из которых состоит из одной материнской цепи и одной (вновь синтезированной) дочерней цепи. *Консервативной единицей, передающейся от одного поколения другому, является одна из двух цепей, составляющих материнский дуплекс.* Таким образом, репликация ДНК носит **полуконсервативный** характер.

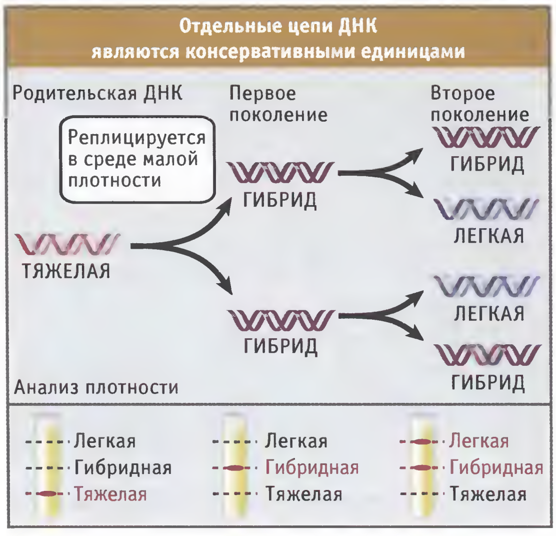


РИС. 1.12 Репликация ДНК полуконсервативна

На рис. 1.12 показан эксперимент, подтверждающий состоятельность данной модели. Если вырастить организм в среде, содержащей подходящий изотоп (например, N^{15}), материнская ДНК будет нести тяжелую метку. При переносе организма в новую среду, содержащую более легкие изотопы, цепи материнской ДНК с тяжелой меткой можно отличить от вновь синтезированных (с легкой меткой) по их массе.

Материнская ДНК состоит из дуплекса двух тяжелых цепей (отмечены красным цветом). В первом поколении, выращенном в легкой среде, ДНК представлена гибридным дуплексом из одной тяжелой материнской цепи (красной) и одной легкой дочерней (синей). Во втором поколении две цепи каждого гибридного дуплекса разъединяются, и каждая из них обретает нового легкого партнера. В результате половина дуплексов остается гибридной, а половина становится полностью легкой (обе цепи голубые).

Вместе с тем отдельные цепи этих дуплексов либо полностью тяжелые, либо полностью легкие. Это утверждение было доказано в 1958 г. экспериментом Мезельсона–Сталя. В эксперименте наблюдали полуконсервативную репликацию ДНК в трех поколениях *E. coli* (как описано выше). Когда из бактерий выделили ДНК и при помощи центрифугирования измерили ее плотность, выяснилось, что ДНК сформировала полосы в соответствии с плотностью: тяжелые для родительских бактерий, гибридные в первом поколении и смесь гибридных и легких во втором поколении.

1.8 Цепи ДНК разделяются в месте, называемом репликативной вилкой

Ключевые понятия:

- Репликацию ДНК контролирует комплекс ферментов, разделяющих материнские цепи и синтезирующих дочерние.
- Репликативная вилка — это точка, в которой разделяются цепи материнской ДНК.
- Ферменты, осуществляющие синтез ДНК, называют ДНК-полимеразами, а ферменты, синтезирующие РНК, называют РНК-полимеразами.
- Разрушение нуклеиновых кислот осуществляют ферменты нуклеазы; нуклеазы принято делить на ДНКазы и РНКазы, обладающие эндо- или экзонуклеазной активностью.

Для того чтобы началась репликация, цепи материнского ДНК-дуплекса должны разъединиться. Однако нарушение двойной спирали носит временный характер. Структура молекулы восстанавливается за счет формирования нового, дочернего дуплекса. В отдельно взятый момент в ДНК расплетен лишь маленький участок двойной спирали.

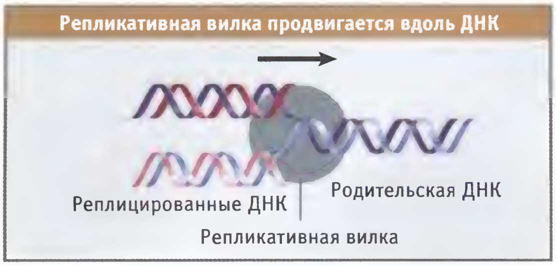


РИС. 1.13 Регион ДНК, в котором расплетенный родительский дуплекс переходит во вновь синтезированные дочерние дуплексы, называется репликативной вилкой

Спиральная структура молекулы ДНК, вовлеченной в репликацию, изображена на РИС. 1.13. Регион молекулы, еще не вовлеченный в репликацию, состоит из материнского дуплекса, а в регионе, где идет репликация, формируется два дочерних дуплекса. Структуру ДНК, возникающую на промежуточном участке между этими регионами, называют **репликативной вилкой**. По мере синтеза дочерних молекул репликативная вилка продвигается вдоль материнской ДНК, в результате чего материнские цепи полностью расплетаются, формируя дочерние дуплексы.

Синтез нуклеиновых кислот катализируется специальными ферментами, распознающими матрицу и обеспечивающими встраивание мономеров в синтезируемую полинуклеотидную цепь. В соответствии с типом вновь синтезируемой цепи ферменты, участвующие в синтезе, разделяют на **ДНК-полимеразы** (синтезируют ДНК) и **РНК-полимеразы** (синтезируют РНК).

Разрушение нуклеиновых кислот также требует участия специальных ферментов: **дезоксирибонуклеаз (ДНКаз)**, которые разрушают ДНК и **рибонуклеаз (РНКаз)**, разрушающих РНК. Нуклеазы принято делить на два основных класса: **эндонуклеазы** и **экзонуклеазы**.

Эндонуклеазы разрушают некоторые связи в молекулах ДНК и РНК, формируя отдельные фрагменты. Некоторые ДНКазы вносят разрыв в обе цепи дуплекса ДНК, а некоторые — лишь в одну из двух цепей. На РИС. 1.14 показано, как эндонуклеазы осуществляют расщепление ДНК.

Экзонуклеазы удаляют основания с концов молекулы одно за другим, приводя к высвобождению мононуклеотидов. Они всегда работают на одной цепи нуклеиновой кислоты и в специфичном для них направлении: начиная с 5'- или 3'-конца. На РИС. 1.15 показан принцип работы экзонуклеаз.

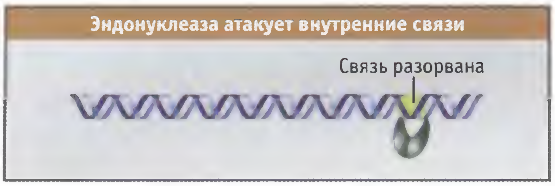


РИС. 1.14 Эндонуклеаза разрезает связь в пределах нуклеиновой кислоты. На рисунке изображен фермент, атакующий одну из цепей ДНК-дуплекса

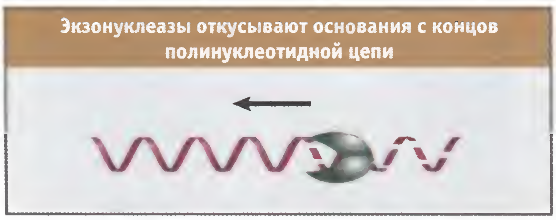


РИС. 1.15 Экзонуклеаза удаляет основания одно за другим путем разрезания последней связи в полинуклеотидной цепи

1.9 Генетическая информация может быть закодирована в ДНК или РНК

Ключевые понятия:

- Большинство геномов представляют собой ДНК, однако геномы вирусов и вироидов могут состоять из РНК.
- Перевод информации из ДНК в РНК происходит в результате транскрипции, а из РНК в ДНК — в ходе обратной транскрипции.
- Трансляция РНК в белок однонаправленна.

Парадигму молекулярной биологии определяет ее **центральная догма**. Генетическая информация записана в последовательностях нуклеиновых кислот в виде генов, а за выполнение закодированных в генах функций отвечают белки. Репликация отвечает за передачу генетической информации по наследству, а транскрипция и трансляция отвечают за перевод этой информации из одной формы в другую.

РИС. 1.16 демонстрирует роль репликации, транскрипции и трансляции с точки зрения центральной догмы молекулярной биологии:

- *Генетический материал сохраняется в нуклеиновых кислотах — ДНК или РНК.* В клетках для хранения генетического материала используется только ДНК. Некоторые вирусы используют для этих целей РНК, а репликация вирусной РНК происходит в инфицированной клетке.
- *Процесс реализации генетической информации обычно однонаправлен.* Транскрипция ДНК приводит к появлению молекул РНК, которые в дальнейшем используются для синтеза аминокислотных последовательностей белков; как

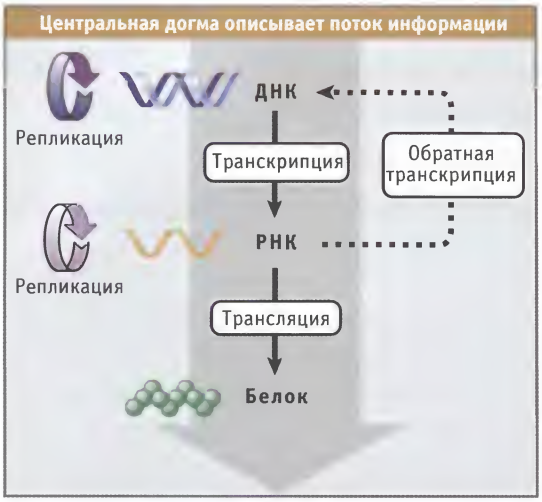


РИС. 1.16 Центральная догма гласит, что информация может сохраняться или переноситься между нуклеиновыми кислотами, однако перенос информации в белок носит необратимый характер

правило, РНК не могут быть использованы в качестве долгосрочного хранилища генетической информации. Трансляция РНК в белки всегда необратима.

Описываемые механизмы обладают одинаковой эффективностью как для генетической информации прокариот и эукариот, так и для информации, содержащейся в вирусах. Геномы всех клеточных

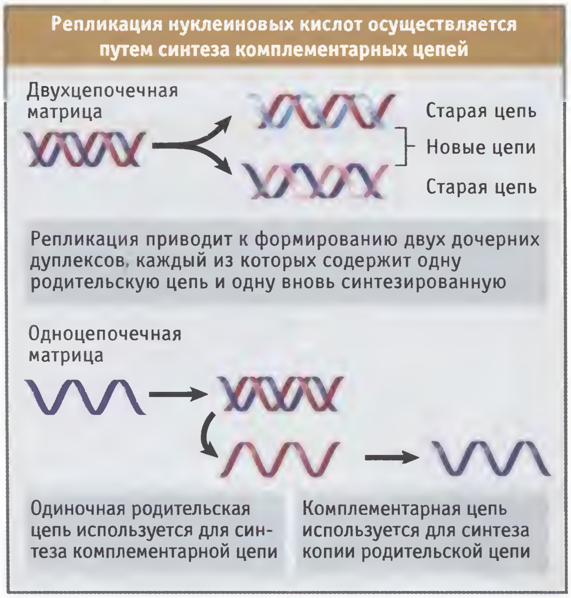


РИС. 1.17 И двухцепочечные, и одноцепочечные нуклеиновые кислоты реплицируются путем синтеза комплементарных цепей, с учетом принципа спаривания оснований

организмов представляют собой ДНК-дуплексы. Геномы вирусов могут состоять из ДНК или РНК, причем независимо от типа нуклеиновой кислоты геном вируса может быть как двух-, так и одноцепочечным. Детали механизма репликации нуклеиновых кислот варьируют в различных вирусных системах, однако все они базируются на принципе комплементарности (РИС. 1.17).

Клетки воспроизводят свою геномную ДНК по механизму полуконсервативной репликации. Двухцепочечные геномы РНК- и ДНК-содержащих вирусов также реплицируются с использованием отдельных цепей дуплекса в качестве матрицы для синтеза комплементарной цепи.

Вирусы, несущие одноцепочечные геномы, для синтеза комплементарной цепи используют свою единственную цепь в качестве матрицы. А уже новая, комплементарная исходному геному, цепь, в свою очередь, используется для синтеза молекул, идентичных оригинальной матрице. В процессе вирусной репликации на некоторых стадиях могут формироваться стабильные или временные двухцепочечные нуклеиновые кислоты.

Ограничение, касающееся однонаправленности переноса информации от ДНК к РНК, не является абсолютным. Оно нарушается **ретровирусами**, чьи геномы состоят из одноцепочечных молекул РНК. В процессе жизненного цикла РНК ретровирусов переводится по механизму **обратной транскрипции** в одноцепочечную молекулу ДНК, а та, в свою очередь, переводится в двухцепочечную ДНК. Этот ДНК-дуплекс становится частью генома зараженной ретровирусом клетки и наследуется как любой другой ген. *Таким образом, обратная транскрипция позволяет восстановить последовательность ДНК на матрице РНК и использовать ее в качестве долгосрочного хранилища генетической информации.*

Существование механизмов обратной транскрипции и репликации РНК устанавливает важный принцип: *информация, содержащаяся в нуклеиновых кислотах любого типа, может быть переведена из одной формы в другую.* В самом распространенном случае клетка использует процессы репликации ДНК, транскрипции и трансляции. Однако в редких случаях (вероятно, опосредованных РНК-содержащими вирусами) информация клеточной РНК переводится в ДНК и встраивается обратно в геном.

Несмотря на то что обратная транскрипция не играет значительной роли во внутриклеточных процессах, она приобретает значимость при рассмотрении вопроса эволюции генома.

Те же принципы сохранения генетической информации справедливы и для больших геномов растений и амфибий, и для сравнительно небольшого генома микоплазмы, и для еще меньших по размеру геномов ДНК- и РНК-содержащих вирусов. На РИС. 1.18 обобщены некоторые примеры, иллюстрирующие разнообразие типов и размеров различных геномов.

Во всем разнообразии организмов, содержащих варьирующие в сотни тысяч раз по размеру геномы, основной принцип остается общим: *ДНК кодирует все белки, которые должны синтезировать клетки организма. Белки прямо или опосредованно выполняют функции, необходимые для выживания организма.* Похожий принцип описывает и суть генетической информации вирусов (вне зависимости от того, содержат они ДНК или РНК): *нуклеиновые кислоты кодируют белки, необходимые для упаковки генома и для выполнения любых других, дополнительных (к выполняемым белками клетки хозяина), функций, служащих цели репродукции вируса на протяжении его инфекционного цикла.* В дополнение следует упомянуть, что некоторые вирусы (например, сателлитный вирус некроза табака STNV) не могут реплицироваться без помощи других вирусов. Так, сателлитному вирусу некроза табака необходимо одновременное присутствие в клетке вспомогательного вируса — вируса некроза табака TNV, — который сам по себе является нормальным инфекционным вирусом.

Геномы сильно различаются по размеру		
Геном	Число генов	Пар оснований
Организмы		
Растения	<50000	$<10^{11}$
Млекопитающие	30000	$\sim 3 \times 10^9$
Черви	14000	$\sim 10^8$
Мухи	12000	$1{,}6 \times 10^8$
Грибы	6000	$1{,}3 \times 10^7$
Бактерии	2000–4000	$<10^7$
Микоплазмы	500	$<10^6$
Вирусы, содержащие двухцепочечную ДНК		
Коровья оспа	<300	187000
Паповавирус (SV40)	~6	5226
Фаг Т4	~200	165000
Вирусы, содержащие одноцепочечную ДНК		
Парововирус	5	5000
Фаг фХ174	11	5387
Вирусы, содержащие двухцепочечную РНК		
Реовирус	22	23000
Вирусы, содержащие одноцепочечную РНК		
Коронавирус	7	20000
Грипп	12	13500
Вирус табачной мозаики (TMV)	4	400
Фаг MS2	4	3569
Сателлитный вирус некроза табака (STNV)	1	1300
Вироиды		
Вироид веретеновидности клубней картофеля (PSTV)	0	359

РИС. 1.18 **Размер составляющих геном нуклеиновых кислот варьирует в широком диапазоне**

1.10 Нуклеиновые кислоты гибридизуются в соответствии с принципом комплементарности

Ключевые понятия:
- Нагревание приводит к разделению цепей ДНК-дуплекса.
- Температура плавления — это температура разделения цепей дуплекса.
- Комплементарные одноцепочечные молекулы ДНК при понижении температуры могут ренатурировать (с образованием дуплекса).
- Процесс денатурации (плавления) и ренатурации (гибридизации) может происходить как между однотипными молекулами, так и между молекулами разного типа. При этом могут формироваться дуплексы ДНК-ДНК, РНК-ДНК и РНК-РНК.
- Способность двух одноцепочечных нуклеиновых кислот гибридизоваться зависит от степени их комплементарности.

Способность разделяться на две цепи без разрушения ковалентных связей является важным свойством двойной спирали. Это, в определенных физиологических условиях, позволяет цепям разъединяться и снова воссоединяться с очень высокой скоростью, что необходимо для обеспечения генетических функций нуклеиновых кислот. Специфичность (точность) процесса денатурации и ренатурации определяется комплементарным спариванием оснований.

Концепция комплементарного спаривания оснований является основной для всех процессов, происходящих с нуклеиновыми кислотами. Ключевым аспектом функционирования двухцепочечной молекулы является возможность временного разрушения пар оснований с сохранением (у образовавшихся одноцепочечных нуклеиновых кислот) способности к формированию дуплекса. На РИС. 1.19 показано, что спаривание оснований позволяет комплементарным одноцепочечным нуклеиновым кислотам формировать дуплексы:

- Внутримолекулярный дуплекс может сформироваться путем спаривания оснований между двумя комплементарными участками одноцепочечной молекулы.
- Одноцепочечные молекулы могут спариваться друг с другом, формируя межмолекулярный дуплекс.

Формирование внутримолекулярных дуплексов чаще всего реализовано для РНК, однако одноцепочечная ДНК также встречается в природе (например, как форма вирусных геномов). Спаривание оснований между двумя независимыми комплементарными одиночными цепями не ограничивается формированием дуплексов ДНК-ДНК или РНК-РНК, а может возникать и между молекулами ДНК и РНК.

Отсутствие ковалентных связей между комплементарными цепями делает удобными манипуляции с ДНК *in vitro*. Нековалентные связи, стабилизирующие двойную спираль, могут быть легко разрушены путем нагревания или помещения ДНК в раствор с

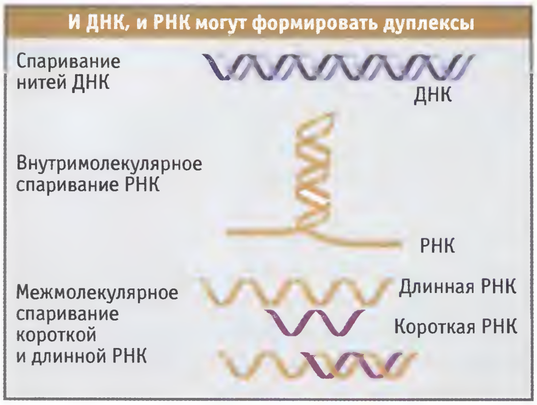


РИС. 1.19 Кроме ДНК дуплексов, основания могут спариваться и при меж- и внутримолекулярных взаимодействиях одноцепочечных РНК (или ДНК)

пониженной концентрацией солей. Если водородные связи между цепями ДНК разрушены, цепи разъединяются.

Процесс разделения цепей носит название **денатурации**, или (чаще употребляемое в разговорной речи) *плавления*. (Термин «денатурация» также используется для описания процесса потери белком своей исходной структуры; это основной термин, отражающий переход естественной конформации макромолекулы в какую-либо иную форму.)

Денатурация ДНК происходит в сравнительно небольшом температурном диапазоне и приводит к серьезным изменениям ее физических свойств. Средняя точка температурного диапазона, в котором происходит разделение цепей ДНК, называется *температурой плавления* (T_m). Она зависит от соотношения в молекуле пар

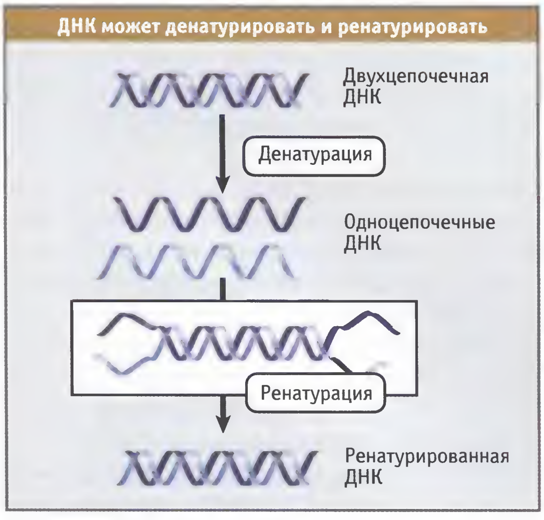


РИС. 1.20 Денатурированные цепи ДНК могут ренатурировать (воссоединяться), вновь образуя дуплекс

G-C и A-T. Поскольку каждую G-C пару связывают три водородные связи, они более стабильны, чем A-T пары, соединенные лишь двумя водородными связями. Чем больше G-C пар содержится в ДНК, тем большее количество энергии необходимо для разделения ее цепей. В растворе с определенными физиологическими условиями ДНК на 40% состоящая из G-C пар (типичное для геномов млекопитающих соотношение) денатурирует при Tm около 87 °C. Таким образом, в диапазоне физиологических температур дуплекс ДНК стабилен.

Денатурация ДНК обратима при соблюдении определенных условий. Способность двух разъединенных комплементарных цепей воссоединяться в двойную спираль носит название **ренатурация**. Процесс ренатурации зависит от специфичного (по последовательности) спаривания оснований в комплементарных цепях. Реакция происходит в две стадии (РИС. 1.20). На первом этапе одиночные цепи ДНК случайно сближаются друг с другом в растворе; если их последовательности комплементарны, две цепи формируют короткий участок двойной спирали. По мере спаривания оснований этот участок увеличивается (этот процесс можно сравнить с застегиванием молнии), в результате формируя длинный молекулярный дуплекс. Ренатурация двойной спирали ДНК восстанавливает исходные свойства молекулы, потерянные во время денатурации.

Ренатурация отражает реакцию между двумя комплементарными последовательностями, которые ранее были разъединены в результате денатурации. Тем не менее дуплекс могут образовать две любые комплементарные молекулы (даже не находившиеся ранее в дуплексе). Иногда это называют **отжигом**. Однако чаще, когда в реакцию вовлечены нуклеиновые кислоты из разных источников, процесс образования дуплекса называют **гибридизацией** (как, например, в том случае, когда один препарат состоит из ДНК, а другой — из РНК). *Способность двух препаратов нуклеиновых кислот гибридизоваться является точным тестом на их комплементарность, поскольку* только *комплементарные последовательности могут формировать дуплексы.*

Методика гибридизации состоит в смешивании двух одноцепочечных нуклеиновых кислот друг с другом (в определенных условиях) и измерении полученного в результате реакции двухцепочечного материала. На РИС. 1.21 изображена процедура, в которой препарат ДНК денатурируют и адсорбируют одиночные цепи нуклеиновой кислоты на специальном фильтре. Затем добавляют второй образец денатурированной ДНК (или РНК). Фильтр обрабатывают таким образом, чтобы второй препарат закрепился на нем только в том случае, если его основания спарились с уже адсорбированной на фильтре ДНК из первого препарата. Обычно второй препарат метят радиоактивной меткой, и после гибридизации можно произвести измерение количества осевшей на фильтре радиоактивной метки.

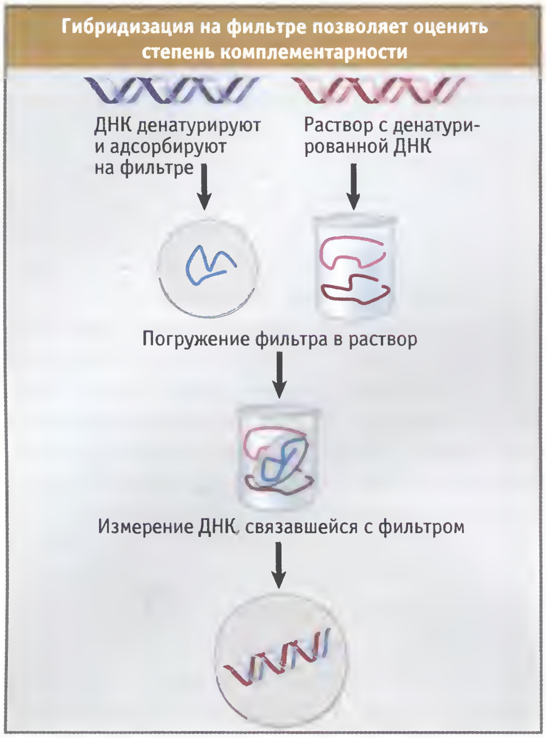


РИС. 1.21. Гибридизация на фильтре позволяет определить, содержит ли раствор денатурированной ДНК (или РНК) последовательности, комплементарные цепям, иммобилизованным на фильтре

Протяженность участка гибридизации между двумя одноцепочечными нуклеиновыми кислотами определяется их комплементарностью. Заметим, что для того, чтобы эффективно гибридизоваться, две последовательности необязательно должны быть полностью комплементарны. Если они не идентичны, но очень похожи, формируется неидеальный дуплекс, в котором в местах несовпадения двух последовательностей формируются неспаренные участки.

1.11 Мутации изменяют последовательность ДНК

Ключевые понятия:

- Все мутации — это изменения в последовательности ДНК.
- Мутации могут возникать спонтанно или могут быть спровоцированы мутагенами.

Мутации наглядно демонстрируют, что ДНК — это генетический материал. Раз изменение в последовательности ДНК приводит к изменению в последовательности белка, значит, эта последовательность ДНК кодирует данный белок. Кроме того, изменения в фенотипе организма позволяют нам определить функцию белка. Если в некотором гене встречается несколько разных мутаций, это может вести к формированию (в популяции) нескольких вариантов белков. Сравнивая и детально анализируя эти варианты, можно выявить регионы белков, ответственные за отдельные ферментативные (или иные) функции.

Все организмы подвержены мутациям, возникающим как в результате внутриклеточных процессов, так и под влиянием окружающей среды. Такие мутации называют **спонтанными**; скорость, с которой они возникают, специфична для каждого отдельно взятого организма и иногда называется **фоновым уровнем** мутирования. Возникновение мутации — довольно редкое событие. Вдобавок, мутации, серьезно нарушающие структуру кодируемого геном белка, элиминируются в процессе эволюции. В этой связи довольно сложно получить для исследований значительное число спонтанных мутантов из естественных популяций.

Частоту возникновения спонтанных мутаций можно повысить, воздействуя на организм определенными веществами. Такие вещества называют **мутагенами**, а изменения, которые они вызывают, — **индуцированными мутациями**. Большинство мутагенов действуют на азотистые основания (изменяя их структуру) или влияют на ДНК путем встраивания в нуклеиновую кислоту дополнительных оснований. Об эффективности мутагена судят по тому, насколько он повышает частоту возникновения мутаций относительно фонового уровня. Используя мутагены, можно спровоцировать множество изменений во всех генах. У бактерий и бактериофагов спонтанные мутации, способные инактивировать функцию гена, возникают с более менее постоянной частотой $3–4 \times 10^{-3}$ на геном за одно поколение. А поскольку размеры генома бактериофагов и бактерий существенно отличаются, это обусловливает различия в частоте возникновения мутаций в пересчете на пару оснований. Такая корреляция позволяет предположить, что общая частота возникновения мутаций находилась под воздействием естественного отбора, который поддерживал баланс между негативным эффектом большинства мутаций и положительным эффектом некоторых из них. Этот вывод подтверждает еще одно наблюдение: у архебактерий, живущих в суровых условиях повышенной температуры и кислотности (когда ДНК повреждается чаще, чем в «обычных» условиях), не отмечается повышения частоты мутаций, а даже наоборот, общая частота находится ниже среднего уровня.

У бактерий частота мутаций составляет 10^{-6} на локус за поколение (или, в среднем, $10^{-9}–10^{-10}$ на пару оснований за поколение). Частота мутирования отдельных пар оснований варьирует в огромном диапазоне (около четырех порядков). Мы не обладаем достаточно точным инструментом измерения частоты возникновения мутаций у эукариот, однако считается, что она примерно равна таковой у бактерий при оценке частоты на локус за поколение (РИС. 1.22).

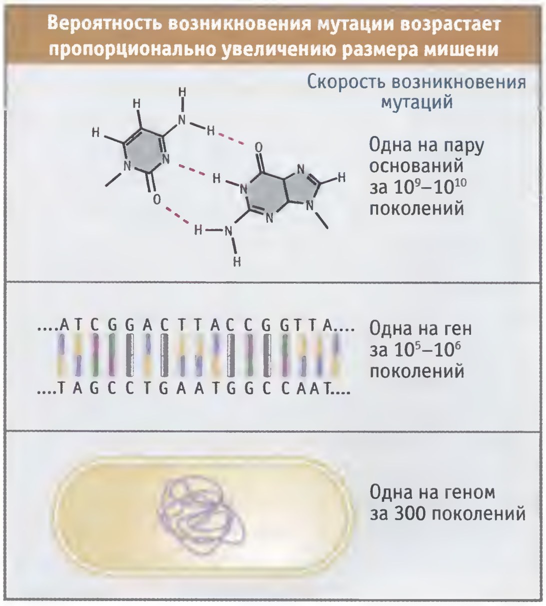


РИС. 1.22 Пара оснований мутирует со скоростью 10^{-9}–10^{-10} на поколение, ген длиной в 1000 пар нуклеотидов (пн) мутирует со скоростью ~10^{-6} на поколение, а бактериальный геном мутирует со скоростью 3×10^{-3} на поколение

1.12 Мутации могут затрагивать отдельные пары оснований или более длинные последовательности

Ключевые понятия:

- Точечные мутации изменяют отдельные пары оснований.
- Точечные мутации могут быть результатом химического превращения одного основания в другое или могут возникать в результате ошибок репликации.
- Замену пары G-C на A-T (или наоборот) называют транзицией.
- Замену пурина на пиримидин (например, A-T на T-A) называют трансверсией.
- Инсерции (вставки) — самый распространенный тип мутаций — происходят в результате перемещения мобильных элементов.

Любая пара оснований в ДНК может мутировать. **Точечные мутации** изменяют только одну пару оснований и могут быть вызваны двумя событиями:

- химической модификацией ДНК, приводящей к превращению одного основания в другое;
- ошибками репликации, приводящими к встраиванию неправильного основания в полинуклеотидную цепь во время синтеза ДНК.

Точечные мутации можно разделить на два класса в зависимости от типа изменяющегося основания.

- Наиболее обширный класс — это **транзиции**, приводящие к замене пиримидина на другой пиримидин или пурина на другой пурин. Они заменяют G-C пары на A-T или наоборот.
- Менее представленный класс — **трансверсии**, в которых пурин заменяется пиримидином или наоборот. Таким образом A-T пара превращается в T-A или C-G пару.

Эффект от воздействия азотистой кислоты представляет собой классический пример транзиции, вызванной химическим превращением одного основания в другое. На РИС. 1.23 показано, что азотистая кислота вызывает окислительное дезаминирование, которое превращает цитозин в урацил. Во время репликации, происходящей после дезаминирования цитозина, U спаривается с A, а не с G (как это было бы без нарушения цитозина). Таким образом, пара C-G заменяется парой T-A (тимин появляется в ДНК на следующем цикле репликации). Азотистая кислота также дезаминирует аденин, приводя к обратной транзиции A-T на G-C.

Кроме того, транзиции могут быть результатом **неправильного спаривания оснований**, когда вопреки модели Уотсона–Крика формируются другие пары. Неправильное спаривание оснований, как правило, является следствием встраивания в ДНК основания с неоднозначными свойствами. На РИС. 1.24 продемонстрирован пример с бромурацилом (BU) — аналогом тимина, содержащим атом брома на месте метильной группы тимина. BU встраивается в ДНК вместо тимина. Однако при спаривании нуклеотидов он обладает неоднозначными свойствами, поскольку наличие брома может приводить к изменению структуры азотистого основания с кето (=O) формы на енольную (–OH) форму. Енольная форма может об-

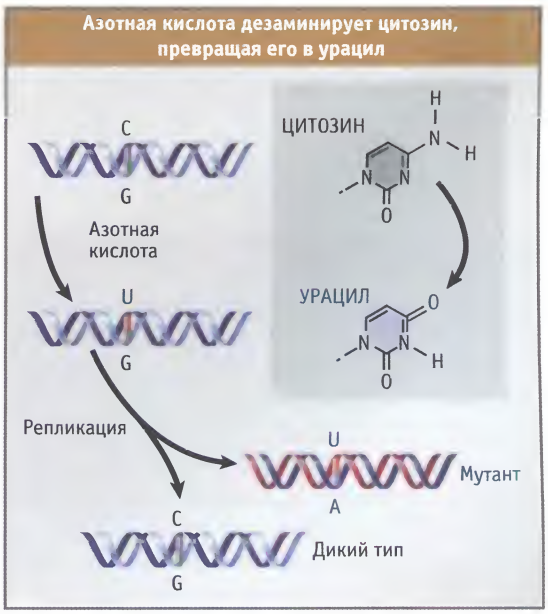


РИС. 1.23 Мутации могут быть вызваны химической модификацией основания

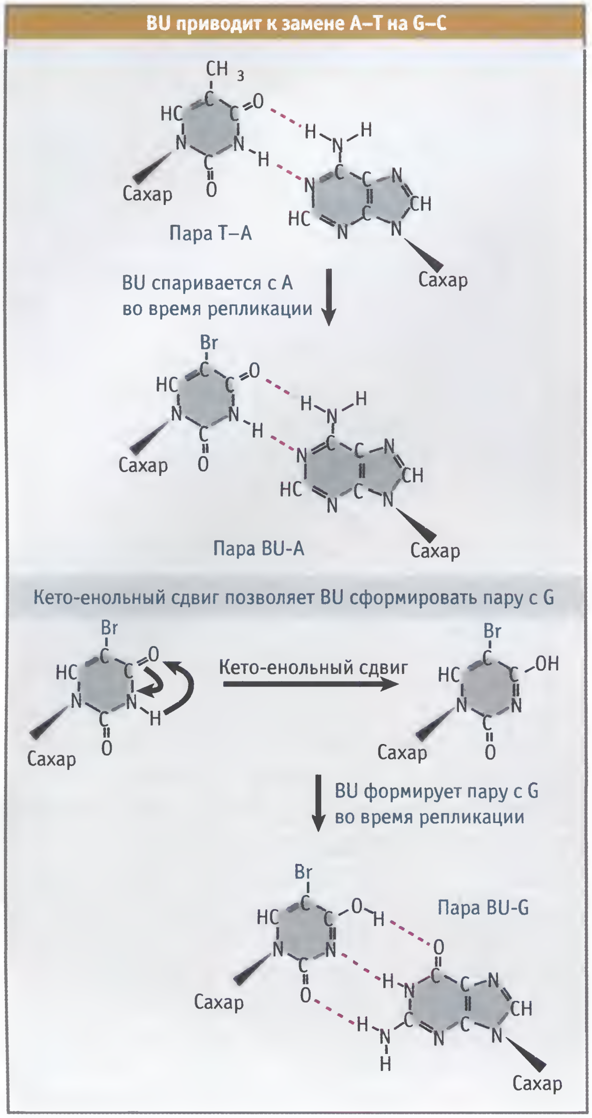


РИС. 1.24 Мутации могут быть вызваны встраиванием в ДНК аналогов оснований

разовывать пару с гуанином, что приводит к замене исходной A-T пары на G-C пару.

Неправильные пары могут появляться как во время исходного встраивания BU, так и во время последующих циклов репликации данной молекулы. Индуцируемая BU транзиция с повышенной вероятностью возникает на каждом цикле репликации однажды «модифицированной» молекулы ДНК. Таким образом, встраивание BU оказывает длительное влияние на последовательность реплицируемых молекул ДНК.

Долгое время было принято считать, что точечные мутации — это основные причины изменения генов. Однако теперь мы знаем, что **инсерции** (вставки) участков дополнительного генетического материала также встречаются довольно часто. Источником «добавляемого» материала чаще всего служат **мобильные элементы**, представляющие собой последовательности ДНК, способные перемещаться из одного места генома в другое (см. гл. 21 «Транспозоны» и гл. 22 «Ретровирусы и ретропозоны»). Инсерции обычно нарушают нормальное функционирование гена. В местах инсерций могут впоследствии возникать **делеции** (удаления) части или всего ранее добавленного материала, а также прилегающих регионов.

Важное отличие точечных мутаций от инсерций/делеций состоит в том, что частота точечных мутаций может быть увеличена с помощью мутагенов, тогда как на частоту возникновения изменений, вызванных мобильными элементами, при помощи мутагенов повлиять нельзя. Однако возникновение инсерций и делеций управляется и другими механизмами (например, ошибками репликации и рекомбинации). По всей видимости, такие вариации встречаются сравнительно редко. Также отметим, что определенный класс мутагенов, называемых акридинами, может приводить к возникновению очень маленьких инсерций и делеций.

1.13 Эффекты мутаций могут быть обратимы

Ключевые понятия:

- Прямые мутации инактивируют ген, а обратные мутации (реверсии) способны компенсировать эти эффекты.
- Эффект от инсерции может быть исправлен делецией вставленного участка, тогда как делеция практически необратима.
- Когда мутация в одном гене сводит на нет эффект от мутации в другом гене, говорят об эффекте подавления (супрессии)

На РИС. 1.25 показано, что возникновение **ревертантов** является важной характеристикой, позволяющей отличать точечные мутации и инсерции от делеций:

- Точечная мутация может быть исправлена в результате восстановления исходной последовательности вследствие возникновения компенсаторной (обратной) мутации.
- Инсерция дополнительного материала может быть исправлена делецией вставленного материала.
- Делеция части гена не может быть исправлена.

Мутации, способные инактивировать ген, называют **прямыми**. Эффект от прямых мутаций может быть компенсирован **обратными** мутациями, которые бывают двух типов: истинные реверсии и вторичные реверсии.

Точное исправление исходной мутации называют **истинной реверсией**. Например, если пара A-T была заменена парой G-C, то мутация, возвращающая A-T пару, приведет к точному восстановлению последовательности исходного (дикого) типа.

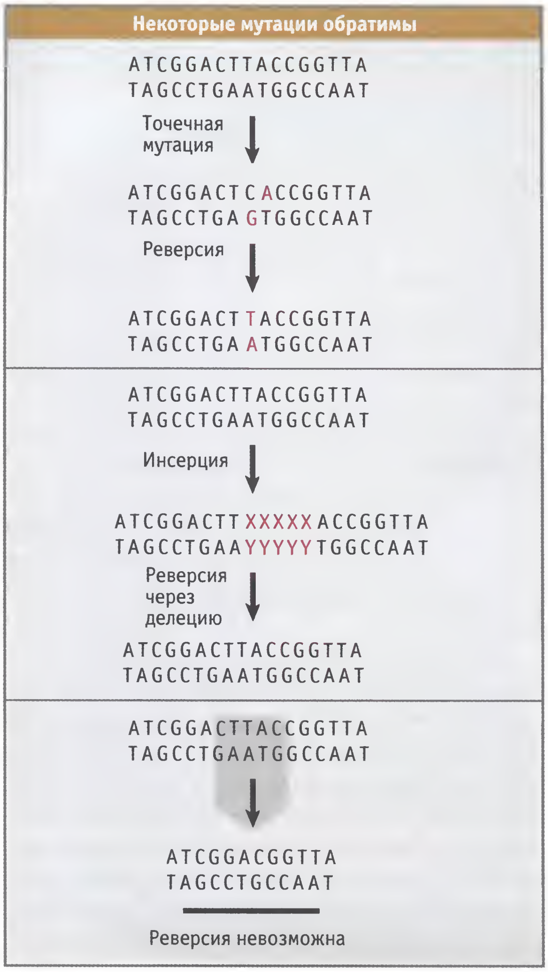


РИС. 1.25 Точечные мутации и инсерции обратимы, а делеции — нет

Второй тип обратных мутаций — так называемые **вторичные реверсии** — могут возникать в другом месте гена, компенсируя при этом эффект от исходной мутации. Например, некая вариация в гене приводит к замене аминокислоты и нарушению свойств белка, а другая мутация компенсирует эффект первой и сохраняет активность белка.

Прямая мутация — это любое изменение, инактивирующее ген, тогда как обратная мутация должна возвратить белку функцию, нарушенную прямой мутацией. Поэтому частота возникновения обратных мутаций значительно ниже, чем прямых (приблизительно на порядок)

Помимо этого, компенсаторный эффект на исходный ген могут оказывать мутации, возникающие в других генах. Этот эффект называют **подавлением (супрессией)**. Локус, мутация в котором компенсирует мутацию другого локуса, называют **супрессором**.

1.14 Мутации наиболее часты в «горячих точках» генома

Ключевые понятия:

- На большей части генома вероятность возникновения мутаций примерно одинакова, однако в «горячих точках» эта вероятность может возрастать на порядки.

До настоящего времени мы рассматривали мутации как индивидуальные изменения в последовательности ДНК, способные влиять на активность генетической единицы, в которой они возникают. Если говорить о мутациях, инактивирующих ген, то частота их возникновения обнаруживает зависимость от размера гена. Следовательно, ген можно рассматривать как мишень для мутаций, и любое его нарушение может вести к нарушению функции. В результате предрасположенность к мутации прямо пропорциональна размеру гена. Но зададимся вопросом: все ли пары оснований одинаково предрасположены к мутациям или некоторые из них мутируют чаще других?

Что получится, если в эксперименте проследить за возникновением мутаций в некотором гене под воздействием мутагенов? Мы получим большое число мутантов, каждый из которых будет результатом индивидуального мутационного явления. Определим местоположение каждой мутации: большинство мутаций будет находиться в разных местах исследуемого гена, однако некоторые места будут подвергаться изменениям чаще других. Две независимо изолированных мутации, расположенные в одном и том же месте ДНК, могут представлять собой либо одну и ту же замену (тогда можно утверждать, что одинаковая мутация возникла более чем в одном случае), либо разные замены (поскольку для каждой пары оснований возможны три различных точечных мутации).

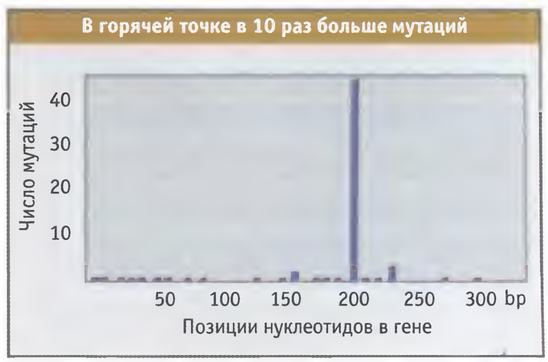


РИС. 1.26 У *E. coli* спонтанные мутации возникают на всем протяжении гена lacI, однако наибольшая их концентрация наблюдается в «горячей точке»

На гистограмме (РИС. 1.26) показаны частоты, с которыми мутации обнаруживаются в каждой паре оснований гена *lacI E. coli.* Статистически вероятность того, что в данном месте возникнет более чем одна мутация, определяется так называемой кинетикой случайных замен (random-hit kinetics) (в соответствии с распределением Пуассона). Таким образом, на некоторых участках будут возникать одна, две или три мутации, тогда как в других их не будет возникать вовсе. В определенных местах частота возникновения мутаций значительно выше, чем ожидаемая частота по случайному распределению: здесь мутации могут возникать в 10 или даже в 100 раз чаще, чем можно ожидать исходя из кинетики случайных замен. Такие места называют **«горячими точками»**. «Горячие точки» могут быть обусловлены как спонтанными мутациями, так и действием различных мутагенов.

1.15 Большинство «горячих точек» появляется в результате модификации оснований

Ключевые понятия:

- Спонтанное дезаминирование, в результате которого происходит превращение 5-метилцитозина в тимин, — одна из наиболее частых причин возникновения «горячих точек».

Основная причина возникновения спонтанных мутаций — это существование необычных оснований в ДНК. Вдобавок к четырем главным основаниям (из которых строится молекула), иногда в ДНК обнаруживают **модифицированные** основания. Название «модифицированные» отражает их природу: они появляются в результате химической модификации одного из четырех оснований, составляющих ДНК. Самое часто встречающееся модифицированное основание — 5-метилцитозин, возникающий за счет добавления метильной группы к некоторым цитозинам в определенных местах последовательности ДНК под действием фермента метилазы.

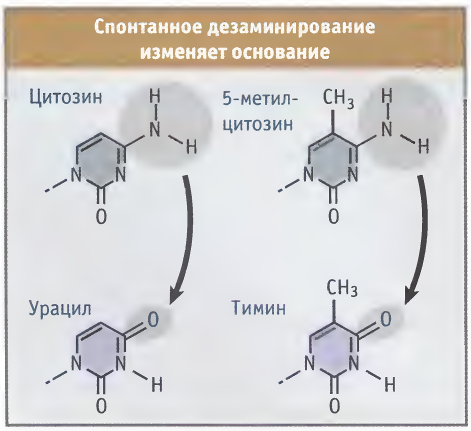


РИС. 1.27 Дезаминирование цитозина приводит к формированию урацила, а дезаминирование 5-метилцитозина — к формированию тимина

Места, содержащие 5-метилцитозин, являются «горячими точками» для спонтанных мутаций у *E. coli.* В каждом таком случае мутация принимает форму транзиции G-C на A-T. У штаммов *E. coli,* не способных метилировать цитозин, отсутствуют «горячие точки» данного типа.

Объяснение существования таких «горячих точек» состоит в том, что цитозины весьма часто подвержены спонтанному дезаминированию. Аминогруппа в этом случае заменяется на кетогруппу. Как было сказано выше, дезаминированный цитозин превращается в урацил (см. рис. 1.23). РИС. 1.27 приводит сравнение этой реакции с дезаминированием 5-метилцитозина, приводящим к формированию тимина. В ДНК это приводит к образованию пар G-U и G-T, где партнеры в паре оснований **не совпадают**.

Все организмы имеют системы репарации, исправляющие неспаренные основания путем удаления и замены. Работа этих систем основана на распознавании неспаренных оснований (типа G-U и G-T), способных приводить к образованию мутаций при репликации.

На РИС. 1.28 показано, что дезаминирование 5-метилцитозина и цитозина приводит к различным последствиям. Дезаминирование весьма редко встречающегося 5-метилцитозина приводит к возникновению мутации, а дезаминирование значительно более часто встречающегося цитозина — нет. Это происходит потому, что репарационные системы распознают пары G-U лучше чем G-T.

У *E.coli* удаление урацила из последовательности ДНК осуществляет фермент урацил-ДНК-гликозилаза (см. разд. 20.5 «Метилазы и гликозилазы вынимают основания из ДНК»). Удаление урацила приводит к появлению неспаренного гуанина, которому репарационная система подбирает комплементарный цитозин. Конечной целью этих реакций служит восстановление исходной последовательности ДНК. Описанная система призвана защищать ДНК от последствий спонтанного дезаминирования цитозина (хотя она и недостаточно активна для того, чтобы полностью предотвратить повышенное дезаминирование, вызванное азотистой кислотой; см. рис. 1.23).

Заметим, что дезаминирование 5-метилцитозина превращает его в тимин, что приводит к формированию участка с неспаренными основаниями G-T. Если эта некомплементарность не будет исправлена до следующего цикла репликации, то после синтеза ДНК в данной позиции возникнет мутация. Во время репликации основания в паре G-T разделяются и образуют пару с новыми «партнерами», приводя к формированию одной пары дикого типа (G-T) и одной мутантной пары (A-T).

Дезаминирование 5-метилцитозина — это одна из самых частых причин возникновения пар G-T в после-

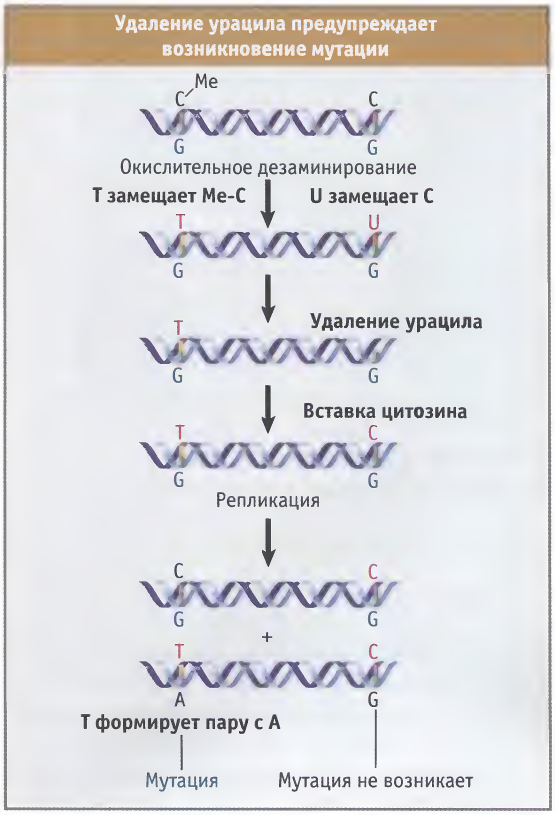


РИС. 1.28 Дезаминирование 5-метилцитозина приводит к формированию тимина (путем замены C-G на T-A), а дезаминирование цитозина приводит к формированию урацила (который обычно удаляется и замещается цитозином)

довательности ДНК. Приоритет репарационных систем, исправляющих неспаренные G-T, сдвинут в сторону замены T на C (а не G на A). Это позволяет снизить частоту возникновения мутаций (см. разд. 20.7 «Контроль направления репарации несовершенных пар»). Однако эти системы не так эффективны, как те, что исправляют пары G-U. Таким образом, дезаминирование 5-метилцитозина чаще приводит к мутациям, чем дезаминирование цитозина.

Отметим также, что 5-метилцитозин дает «горячие точки» в ДНК эукариот. Они часто встречаются в местах концентрации CpG динуклеотидов (так называемых CpG островках (см. разд. 24.19 «CpG островки являются мишенями для регуляции»)). Несмотря на то что в ДНК человека на долю 5-метилцитозина приходится около 1% оснований, доля участков, содержащих эти модифицированные основания, составляет около 30% от всех точечных мутаций, что делает 5-метилцитозин ключевым фактором возникновения мутаций в клетках животных.

Роль репарационных систем как механизма снижения частоты возникновения мутаций можно оценить, наблюдая последствия блокирования мышиного фермента MBD4. Этот фермент представляет собой гликозилазу, способную удалять T (или U), стоящий напротив гуанина. В результате инактиваций фермента частота возникновения мутаций на CpG участках возрастает в три раза. (Эффект не столь значителен, как можно было бы ожидать, поскольку MBD4 — лишь один из представителей системы ферментов, способных исправлять G-T пары. Можно предположить, что нарушение всей системы ферментов привело бы к значительно большему увеличению частоты возникновения мутаций.)

Работа описанной выше системы позволяет по-новому взглянуть на использование в ДНК именно тимина (а не урацила, как в РНК). Возможно, это связано с повышенным требованием к молекуле ДНК в части стабильности ее последовательности. Использование T предполагает, что любые дезаминирования C немедленно распознаются, поскольку приводят к формированию нестандартного для ДНК основания (U), что существенно увеличивает эффективность работы репарационных систем (по сравнению с распознаванием пар G-T, в которых оба основания являются нормальными для ДНК). Кроме того, фосфодиэфирная связь в паре G-U более подвижна.

1.16 Некоторые элементы наследственности чрезвычайно малы

Ключевые понятия:

- Некоторые очень маленькие элементы наследственности не кодируют собственных белков, однако, представляя собой РНК или белок, обладают наследственным потенциалом.

Вироиды — это инфекционные факторы, вызывающие заболевания высших растений и представляющие собой небольшие молекулы РНК. В отличие от вирусов, у которых инфекционный агент представлен **вирионом** (состоящим из заключенного в белковую оболочку генома), *вироидная РНК сама по себе является инфекционным агентом*. Вироид состоит только из РНК, образующей двухцепочечную структуру с небольшими неспаренными участками (РИС. 1.29). Мутации, нарушающие структуру РНК, уменьшают инфекционные свойства вироида.

Вироидная РНК состоит из однотипных молекул, способных автономно реплицироваться в инфицированных клетках растения. Ее последовательность точно воспроизводится в поколениях вироида. Вироиды разделяют на несколько групп по последовательности их РНК. Вироид относят в ту или иную группу по степени сходства его РНК с РНК остальных представителей группы. Например, четыре вироида имеют от 70 до 83% сходства по последовательности РНК с родственным им PSTV (вироид веретеновидности клубней картофеля). Различные изоляты определенного штамма вироидов отличаются

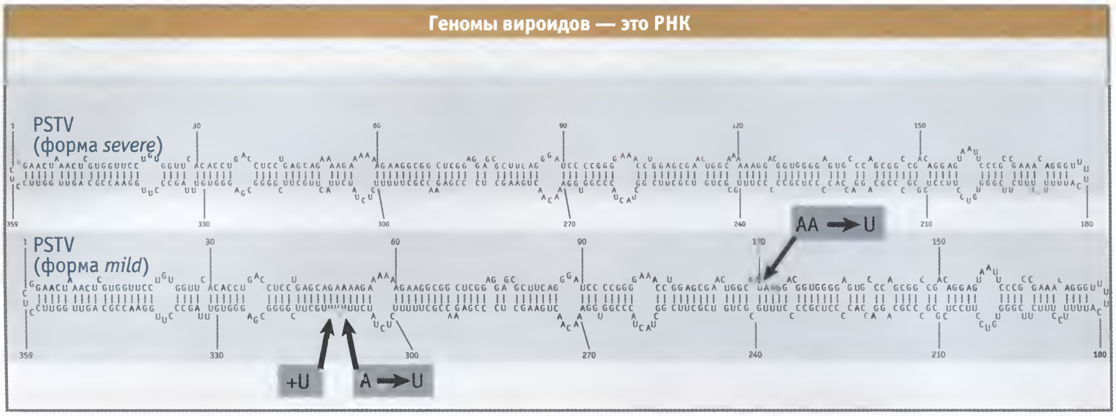


РИС. 1.29 РНК PSTV представляет собой кольцевую молекулу, формирующую протяженную двухцепочечную структуру, прерываемую множеством внутренних петель. Формы *severe* и *mild* различаются по трем сайтам.

друг от друга, и эти отличия могут оказывать влияние на фенотип инфицированной клетки. Например, штаммы *mild* и *severe* PSTV отличаются тремя нуклеотидными заменами.

Вироиды напоминают вирусы, поскольку имеют небольшие геномы, состоящие из нуклеиновых кислот, и удовлетворяют критериям наследования генетической информации. Однако вироиды, которые также иногда называют **субвирусными патогенами**, отличаются от вирусов и по структуре, и по функции. РНК вироида не транслируется в белки и, таким образом, как бы не несет информации, способствующей его выживанию. В связи с этим возникает два вопроса: как реплицируется вироидная РНК и как она влияет на фенотип инфицированной клетки?

Репликация, по всей видимости, реализуется за счет ферментов клетки хозяина. Способность последовательности вироидной РНК передаваться по наследству свидетельствует о том, что РНК выступает в роли матрицы.

Вироиды условно патогенны, поскольку их вмешательство в нормальные клеточные процессы носит случайный характер. Так, например, они могут задействовать важные клеточные ферменты для целей собственной репликации или вмешиваться в процесс производства необходимых клетке РНК. С другой стороны, они могут вести себя как отклоняющиеся от нормы регуляторные молекулы, осуществляющие определенные функции в клетке, вплоть до регуляции экспрессии отдельных генов.

Еще более необычный фактор — это **скрэпи** — причина нейродегенеративного заболевания коз и овец. Данное заболевание родственно заболеванию куру и синдрому Кройцфельда–Якоба, нарушающим функционирование головного мозга у людей.

Скрэпи опосредовано не нуклеиновыми кислотами. Удивительный фактор, вызывающий скрэпи, был назван **прионом** (белковый инфекционный агент, proteinaceous infectious agent). По химической структуре — это гидрофобный гликопротеин **PrP** массой 28 кДа. PrP кодируется ядерными генами млекопитающих, в норме экспрессирующимися в мозге. Этот белок представлен двумя формами: первая называется PrP^c и представляет собой продукт, в норме встречающийся в мозге и эффективно разрушаемый протеазами; вторая называется PrP^{sc} и, в отличие от первой формы, является чрезвычайно устойчивым к разрушающему действию протеаз белком. PrP^c переходит в форму PrP^{sc} путем недостаточно изученного механизма модификации или конформационного изменения, делающего белок устойчивым к действию протеаз.

Инфекционный агент, вызывающий скрэпи, должен каким-то образом изменять механизм синтеза нормального белка так, чтобы он становился устойчив к протеазам (см. разд. 31.12 «Прионы как причина заболеваний млекопитающих»). Мыши, дефицитные по гену PrP, не подвержены скрэпи, что демонстрирует ключевую роль белка PrP в развитии заболевания.

1.17 Резюме

Утверждение о том, что ДНК является генетическим материалом, доказывают два эксперимента. ДНК, изолированная из одного штамма бактерии *Pneumococcus*, способна передавать свойства этого штамма другому штамму. Кроме того, ДНК — это единственный компонент, передаваемый по наследству от материнских фагов к дочерним. ДНК может быть использована для придания новых свойств эукариотическим клеткам.

ДНК представляет собой двойную спираль, состоящую из антипараллельных цепей, составные части которых — нуклеотиды — соединены 5′-3′-фосфоди-

эфирными связями. Каркас молекулы образует сахарофосфатный остов; пуриновые и пиримидиновые основания соединены парами, в которых A комплементарен T, а G комплементарен C. В реакции полуконсервативной репликации цепи ДНК способны разъединяться и, в соответствии с правилом комплементарности, достраиваться с образованием дочерних цепей. Правило комплементарности оснований также выполняется и при транскрипции РНК, которая соответствует одной из цепей ДНК-дуплекса.

Участок ДНК может кодировать белок. Генетический код определяет связь между последовательностью ДНК и последовательностью белка. Только одна из двух цепей ДНК кодирует последовательность аминокислот белка. Кодон состоит из трех нуклеотидов, соответствующих одной аминокислоте. Кодирующая белок последовательность ДНК состоит из серии кодонов, считываемых от определенной точки. Обычно лишь одна из трех возможных рамок считывания может быть транслирована в белок.

Мутация представляет собой изменение в последовательности A-T и G-C пар ДНК. Мутация в кодирующей последовательности может изменить последовательность аминокислот в соответствующем белке. Вставка или делеция нуклеотида может приводить к сбою рамки считывания, изменяя последующие кодоны; это приводит к формированию серии новых аминокислот, закодированных на участке ДНК, следующем за мутацией. Точечные мутации изменяют единственную аминокислоту (кодируемую мутировавшим кодоном). Обратные мутации приводят к исправлению ранее возникшей мутации и восстановлению изначальной последовательности ДНК. Эффект от инсерции может быть обращен вследствие потери ранее вставленного в ДНК дополнительного материала, однако делеции необратимы. Мутации могут быть компенсированы опосредованно, например когда мутация в одном гене сводит на нет эффект от мутации в другом гене.

Частота возникновения мутаций может быть увеличена мутагенами. Мутации чаще возникают в «горячих точках». Чаще всего «горячая точка» является следствием дезаминирования модифицированного основания — 5-метилцитозина.

Прямые мутации возникают с частотой 10^{-6} на локус на поколение; обратные мутации встречаются существенно реже. Не все мутации оказывают влияние на фенотип.

Несмотря на то что вся генетическая информация клеточных организмов представлена двухцепочечной ДНК, у вирусов геномы могут быть представлены двухцепочечной или одноцепочечной ДНК или РНК. Вирусные патогены — вироиды — состоят из небольших кольцевых молекул РНК, не покрытых защитной оболочкой. Их РНК не кодирует белков, а способ патогенеза и сохранения генетического материала для них неизвестен. Скрэпи представляет собой белковый инфекционный фактор.

Литература

1.1 Введение

Обзор

Cairns, J., Stent, G., and Watson, J. D. (1966). Phage and the Origins of Molecular Biology. *Cold Spring Harbor Symp. Quant. Biol.*

Judson, H. (1978). *The Eighth Day of Creation.* Knopf, New York.

Olby, R. (1974). *The Path to the Double Helix.* MacMillan, London.

1.2 ДНК — генетический материал бактерий

Исследование

Avery, O. T., MacLeod, C. M., and McCarty, M. (1944). Studies on the chemical nature of the substance inducing transformation of pneumococcal types. *J. Exp. Med.* 98, 451–460.

Griffith, F. (1928). The significance of pneumococcal types. *J. Hyg.* 27, 113–159.

1.3 ДНК — генетический материал вирусов

Исследование

Hershey, A. D. and Chase, M. (1952). Independent functions of viral protein and nucleic acid in growth of bacteriophage *J. Gen. Physiol.* 36, 39–56.

1.4 ДНК — генетический материал клеток животных

Исследование

Pellicer, A., Wigler, M., Axel, R., and Silverstein, S. (1978). The transfer and stable integration of the HSV thymidine kinase gene into mouse cells. *Cell* 14, 133–141.

1.6 Структура ДНК представлена двойной спиралью

Исследование

Watson, J. D., and Crick, F. H. C. (1953). A structure for DNA. *Nature* 171, 737–738.

Watson, J. D., and Crick, F. H. C. (1953). Genetic implications of the structure of DNA. *Nature* 171, 964–967.

Wilkins, M. F. H., Stokes, A. R., and Wilson, H. R. (1953). Molecular structure of DNA. *Nature* 171, 738–740.

1.7 Процесс репликации ДНК полуконсервативен

Обзор

Holmes, F. (2001). *Meselson, Stahl, and the Replication of DNA: A History of the Most Beautiful Experiment in Biology.* Yale University Press, New Haven, CT.

Исследование

Meselson, M. and Stahl, F. W. (1958). The replication of DNA in *E. coli. Proc. Natl. Acad. Sci. USA* 44, 671–682.

1.11 Мутации изменяют последовательность ДНК

Обзор

Drake, J. W., Charlesworth, B., Charlesworth, D., and Crow, J. F. (1998). Rates of spontaneous mutation. *Genetics* 148, 1667–1686.

Drake, J. W. and Balz, R. H. (1976). The biochemistry of mutagenesis. *Annu. Rev. Biochem.* 45, 11–37.

Исследование

Drake, J. W. (1991). A constant rate of spontaneous mutation in DNA-based microbes. *Proc. Natl. Acad. Sci. USA* 88, 7160–7164.

Grogan, D. W., Carver, G. T., and Drake, J. W. (2001). Genetic fidelity under harsh conditions: analysis of spontaneous mutation in the thermoacidophilic archaeon *Sulfolobus acidocaldarius. Proc. Natl. Acad. Sci. USA* 98, 7928–7933.

1.12 Мутации могут затрагивать отдельные пары оснований или более длинные последовательности

Обзор

Maki, H. (2002). Origins of spontaneous mutations: specificity and directionality of base-substitution, frameshift, and sequence-substitution mutageneses. *Annu. Rev. Genet.* 36, 279–303.

1.15 Большинство «горячих точек» мутагенеза появляется в результате модификации оснований

Исследование

Coulondre, C. et al. (1978). Molecular basis of base substitution hotspots in *E. coli. Nature* 274, 775–780.

Millar, C. B., Guy, J., Sansom, O. J., Selfridge, J., MacDougall, E., Hendrich, B., Keightley, P. D., Bishop, S. M., Clarke, A. R., and Bird, A. (2002). Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. *Science* 297, 403–405.

1.16 Некоторые элементы наследственности чрезвычайно малы

Обзор

Diener, T. O. (1986). Viroid processing: a model involving the central conserved region and hairpin. *Proc. Natl. Acad. Sci. USA* 83, 58–62.

Diener, T. O. (1999). Viroids and the nature of viroid diseases. *Arch. Virol. Suppl.* 15, 203–220.

Prusiner, S. B. (1998). Prions. *Proc. Natl. Acad. Sci. USA* 95, 13363–13383.

Исследование

Bueler, H. et al. (1993). Mice devoid of PrP are resistant to scrapie. *Cell* 73, 1339–1347.

McKinley, M. P., Bolton, D. C., and Prusiner, S. B. (1983). A protease-resistant protein is a structural component of the scrapie prion. *Cell* 35, 57–62.

Гены кодируют белки

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

2.1 Введение

2.2 Один ген кодирует один полипептид

- Гипотеза «один ген — один фермент» составляет основу современной генетики: ген — это участок ДНК, кодирующий одну полипептидную цепь.
- Мутация может нарушить функцию гена.

2.3 Мутации в одном гене не могут дополнять друг друга

- Мутация в конкретном варианте гена (аллеле) влияет только на белок, закодированный данным аллелем, и не влияет на варианты белка, закодированные другими аллелями гена.
- Если две мутации в гетерозиготном состоянии не компенсируют друг друга (не формируют дикий фенотип), это означает, что они являются частью одного гена.

2.4 Мутации могут оказывать как положительный, так и отрицательный эффект

- Рецессивные мутации представляют собой нарушения функции белка.
- Доминантные мутации ведут к усилению функции белка.
- Определение взаимосвязи гена и фенотипа возможно через возникновение нуль-мутации (мутации, полностью элиминирующей функцию гена).
- Молчащие мутации не оказывают эффекта на фенотип, поскольку изменение нуклеотидного состава не изменяет последовательность или количество белка либо, поскольку изменение аминокислотной последовательности не оказывает эффекта на функцию белка.
- Некоторые мутации влияют на функцию белка, однако практически не влияют на фенотип, поскольку белок сохраняет достаточную активность.

2.5 Локус может иметь различные мутантные аллели

- Существование множества аллелей определяет существование гетерозигот, содержащих любую комбинацию двух аллелей.

2.6 Локус может иметь более одного аллеля дикого типа

- Локус может быть представлен несколькими часто встречающимися аллелями, в этом случае среди них сложно выделить единственный дикого типа.

2.7 Рекомбинация — это результат физических перестроек ДНК

- Рекомбинация — это результат процесса, называемого кроссинговером; он происходит с образованием специальных структур — хиазм и вовлекает две из четырех хроматид.
- Процесс рекомбинации включает разрыв и воссоединение цепей ДНК с формированием участков гибридной ДНК.

2.8 Генетический код триплетен

- Генетический код считывается триплетами нуклеотидов (кодонами).
- Триплеты не перекрываются и начинают считываться с определенной стартовой точки.
- Мутации, представляющие собой вставку или удаление отдельных нуклеотидов, приводят к сбою в считывании триплетов на участке ДНК после мутации.
- Сочетание мутаций, приводящих к добавлению или удалению сразу трех (или кратного трем числа) нуклеотидов, добавляет или убирает одну (несколько) аминокислот в белке, но не меняет рамку считывания триплетов после места расположения мутации.

2.9 Существует три возможных рамки считывания каждой последовательности

- Как правило, транслируется только одна из трех рамок считывания, остальные заблокированы сигналами терминации трансляции.

2.10 Прокариотические гены колинеарны их белкам

- Прокариотический ген включает участок из $3N$ нуклеотидов, кодирующих N аминокислот.
- Ген, мРНК и белок колинеарны друг другу.

2.11 Синтезу белка предшествует ряд процессов

- Прокариотический ген экспрессируется через транскрипцию (синтез мРНК) и трансляцию (мРНК) с образованием белка.
- У эукариот ген может содержать внутренние регионы, не представленные в белке.
- Внутренние регионы удаляются из эукариотического РНК-транскрипта посредством РНК-сплайсинга. В итоге эукариотическая мРНК также оказывается колинеарна белку.
- Молекула мРНК содержит кодирующий регион и нетранслируемые 5′-регион и 3′-регион.

2.12 Регуляторные белки — *транс*-действующие, а регуляторные участки ДНК — *цис*-действующие

- РНК и белки являются *транс*-действующими агентами, то есть могут воздействовать на любую из копий регулируемого ими гена в данной клетке.
- *Цис*-действующие мутации позволяют определять последовательности ДНК, являющиеся мишенями для *транс*-действующих РНК или белков. В отличие от РНК или белков *цис*-действующие мутации не перемещаются в пространстве и могут оказывать влияние только на данную молекулу ДНК.

2.13 Резюме

2.1 Введение

Ген — это функциональная единица наследственности. Каждый ген — это последовательность нуклеотидов, несущая информацию о конкретном продукте (РНК или белке). Ключевые свойства гена были установлены еще Менделем более века назад. Эти свойства, сформулированные им в двух законах, представляют ген как дискретный фактор, передаваемый в неизменном виде от родителей к потомству. Ген может быть представлен несколькими альтернативными вариантами, которые называют аллелями.

В диплоидных организмах, имеющих в клетке по два набора хромосом, от каждого из родителей потомству передается по одной копии каждой хромосомы, следовательно, и записанных в ней генов. Одна из двух находящихся в клетке копий гена является аллелем, унаследованным от отца, другая — унаследованным от матери. Сходство в поведении генов и хромосом позволило исследователям предположить, что гены расположены на хромосомах.

Хромосома содержит линейно расположенные на ней гены, причем каждый ген занимает на хромосоме строго определенное положение. Местоположение гена формально принято называть генетическим локусом, а аллели гена — это слегка отличающиеся варианты последовательности, находящиеся в данном генетическом локусе.

Ключом к формированию понятий об организации генов в хромосомах послужило открытие сцепленности генов — тенденции некоторых генов (как позже выяснилось, находящихся на одной хромосоме) оставаться связанными при исследовании, хотя в соответствии с законами Менделя они должны были комбинироваться независимо. Поскольку единица рекомбинации (перемешивания) генетической информации была выражена через значение сцепленности, стало возможным создание генетических карт (расположения генов на хромосомах).

Разрешающая способность карт рекомбинации высших эукариот ограничена небольшим числом потомков, которых можно получить от одного скрещивания. Между близко расположенными местами хромосомы рекомбинация происходит крайне редко (например, используя в качестве генетических маркеров две мутации в одном гене, наблюдать рекомбинацию невозможно). Таким образом, классические карты сцепления для хромосом эукариот помогают расставлять гены по порядку, но не могут отражать взаимосвязи в пределах гена. Возникновение рекомбинации в пределах одного гена можно продемонстрировать в бактериальных системах, где каждое скрещивание приводит к появлению гораздо большего числа потомков.

Мутации в гене также организованы в линейном порядке, демонстрируя, что сам по себе ген имеет такую же линейную структуру, что и совокупность генов на хромосоме. Таким образом, генетическая карта хромосомы линейна (как в пределах локуса, так и между локусами); она состоит из непрерывной последовательности нуклеотидов, на которой и расположены гены. Этот вывод соответствует современным представлениям об устройстве генетической информации, в общем виде приведенным на РИС. 2.1, — генетический материал хромосомы состоит из молекулы ДНК, на которой один за другим расположены гены.

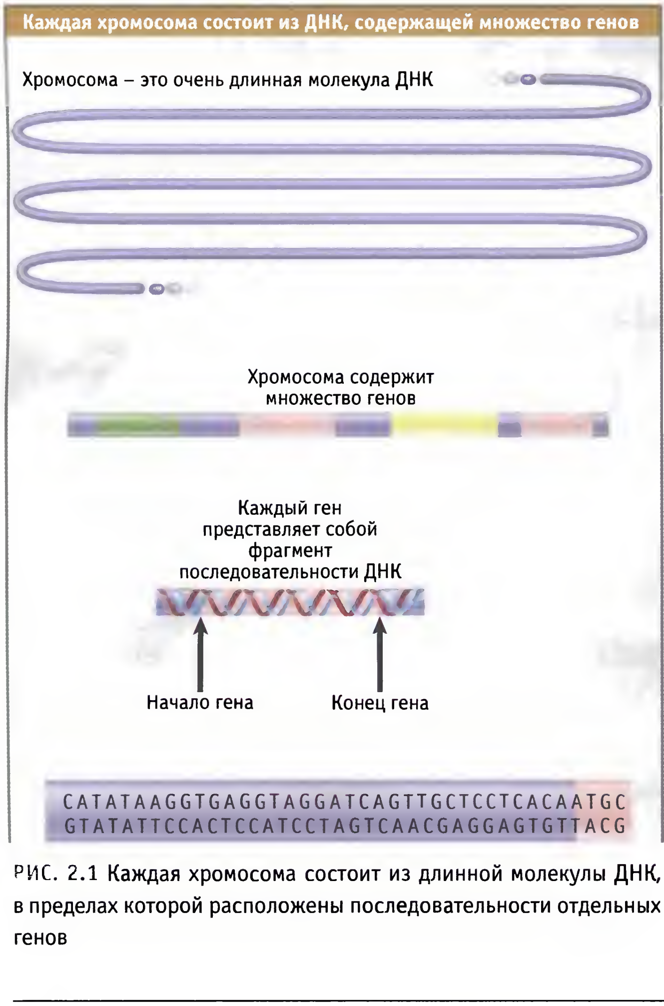


РИС. 2.1 Каждая хромосома состоит из длинной молекулы ДНК, в пределах которой расположены последовательности отдельных генов

2.2 Один ген кодирует один полипептид

Ключевые понятия:

- Гипотеза «один ген — один фермент» составляет основу современной генетики: ген — это участок ДНК, кодирующий одну полипептидную цепь.
- Мутация может нарушить функцию гена.

Первые попытки ассоциировать гены с ферментами показали, что катализируемая ферментом стадия некоторого метаболического пути может быть блокирована мутацией в гене этого фермента. Так сформировалась гипотеза *один ген — один фермент*. Каждый шаг метаболизма катализируется определенным ферментом, за продукцию

которого ответственен один ген, и мутация в этом гене влияет на активность закодированного фермента.

Эта гипотеза требует некоторого уточнения для ферментов, состоящих более чем из одной субъединицы. Если все субъединицы одинаковы, то такой фермент — **гомомультимер** — действительно кодируется одним геном. Если же субъединицы различаются, фермент называется **гетеромультимером**. Для гетеромультимеров гипотеза «один ген — один фермент» является недостаточно общей, точнее будет сказать *один ген — одна полипептидная цепь.*

Довольно сложно бывает определить, какой именно белок кодирует данный ген. Мутация, которая в опытах Менделя приводила к формированию гороха с морщинистым фенотипом, была описана как вариация, инактивирующая ген фермента разветвления крахмала, лишь в 1990 году!

Важно помнить, что ген формирует белок не напрямую. Как было показано на рис. 1.2, с гена считывается РНК, которая, в свою очередь, участвует в синтезе белка. Большинство генов несут информацию о строении белков, однако некоторые из них кодируют РНК, не связанные с синтезом белков. Эти РНК могут быть структурными компонентами аппарата, ответственного за синтез белков, или могут играть роль в процессе регуляции экспрессии генов. Основной принцип состоит в том, что ген — это последовательность ДНК, которая определяет структуру некоторого независимого продукта. Процесс экспрессии гена может завершиться формированием конечного продукта как в виде белка, так и в виде РНК.



РИС. 2.2 Гены кодируют белки; доминирование объясняется свойствами мутантных белков. Рецессивный аллель не вносит вклада в формирование фенотипа, поскольку он не участвует в формировании белка (или участвует, но синтезируемый белок является нефункциональным)

Мутация — это случайное событие, связанное с изменением структуры гена. Следовательно, вероятность того, что мутация нарушит или даже аннулирует функцию гена, достаточно высока. Большинство мутаций, влияющих на функцию генов, рецессивны. *Они связаны с нарушением кодируемой геном функции (мутантный ген не может обеспечить наработку нормального белка).* РИС. 2.2 иллюстрирует взаимодействие рецессивных аллелей и аллелей дикого типа. Когда клетка содержит один аллель дикого типа и один мутантный аллель (гетерозиготное состояние), аллель дикого типа способен обеспечить продукцию нормального фермента. В этом случае говорят, что аллель дикого типа является доминантным (т. е. единственный аллель дикого типа обеспечивает наработку достаточного количества активного белка). Если же количество белка, производимое единственным функциональным аллелем, недостаточно для сохранения дикого фенотипа, говорят о формировании промежуточного фенотипа (и частично доминантном аллеле в гетерозиготе).

2.3 Мутации в одном гене не могут дополнять друг друга

Ключевые понятия:

- Мутация в конкретном варианте гена (аллеле) влияет только на белок, закодированный данным аллелем, и не влияет на варианты белка, закодированные другими аллелями гена.
- Если две мутации в гетерозиготном состоянии не компенсируют друг друга (не формируют дикий фенотип), это означает, что они являются частью одного гена.

Как определить, лежат ли две мутации, приводящие к формированию одного и того же фенотипа, в пределах одного гена? Если генетические карты мутантных генов схожи, они могут быть аллелями. Однако эти мутации могут находиться и в двух *разных* генах, продукты которых выполняют одинаковую функцию. Для того чтобы понять, находятся мутации в одном гене или в разных, используют **комплементационный тест**. Тест заключается в том, что получают гетерозиготное по обеим мутациям поколение организмов (для этого скрещивают родителей, гомозиготных по каждой мутации).

Если мутации находятся в одном гене, генотипы гомозиготных родителей будут выглядеть следующим образом:

$$\frac{m_1}{m_1} \quad и \quad \frac{m_2}{m_2}$$

Первый родитель передает по наследству один мутантный аллель (m_1), а второй родитель — другой (m_2), таким образом, гетерозиготу можно записать как:

$$\frac{m_1}{m_2}$$

Поскольку ген дикого типа у потомка отсутствует, он будет иметь мутантный фенотип.

Если же мутации находятся в разных генах, то генотипы родителей можно записать так:

$$\frac{m_1 +}{m_1 +} \quad \text{и} \quad \frac{+ m_2}{+ m_2}$$

Каждая хромосома в таком случае будет иметь одну копию гена дикого типа (обозначенную знаком плюс) и вторую — мутантную — копию. Генотип потомка от скрещивания приведенных выше родителей можно записать как:

$$\frac{m_1 +}{+ m_2}$$

В данном случае гетерозиготный потомок унаследует по одной функциональной копии каждого гена (+) и будет иметь дикий фенотип. Принято говорить, что такие гены **комплементарны**.

Более подробно тест на комплементацию показан на РИС. 2.3. Простейший тест представляет собой сравнение, изображенное в верхней части рисунка. Если две мутации находятся в одном и том же гене, мы видим различия в фенотипах *транс-* и *цис*-конфигурации. *Транс*-конфигурация мутантна, поскольку каждый аллель несет мутацию. Однако *цис*-конфигурация дает дикий фенотип, поскольку один аллель имеет две мутации, а другой аллель мутаций не несет. В нижней части рисунка показано, что если две мутации находятся в разных генах, формируемый фенотип всегда будет дикого типа. В любом случае в клетке окажется один аллель дикого типа и один мутантный аллель для каждого из двух генов. Если же тест на комплементацию не проходит, значит, две мутации являются частью *одной* генетической единицы. Принято говорить, что мутации, которые не дополняют друг друга, составляют часть одной **комплементационной группы**. Другой термин, используемый для описания единицы генетической информации, определенной тестом на комплементацию — **цистрон**. По сути это то же, что и ген. Все три термина описывают участок ДНК, действующий как единица, способная производить продукт в виде РНК или белка. Свойства гена в части комплементации объясняются тем, что продукт гена является молекулой, ведущей себя как независимая функциональная единица.

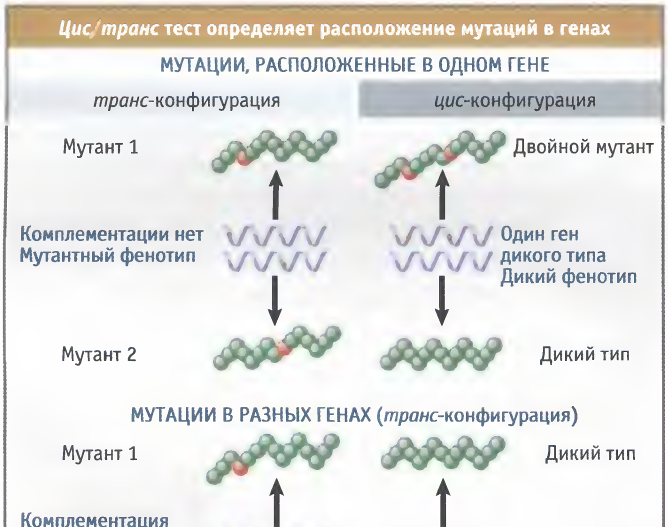


РИС. 2.3 Для определения цистрона используют комплементационный тест. Гены представлены в виде спиралей. Красные звездочки определяют место расположения мутаций

2.4 Мутации могут оказывать как положительный, так и отрицательный эффект

Ключевые понятия:

- Рецессивные мутации представляют собой нарушения функции белка.
- Доминантные мутации ведут к усилению функции белка.
- Определение взаимосвязи гена и фенотипа возможно через возникновение нуль-мутации (мутации, полностью элиминирующей функцию гена).
- Молчащие мутации не оказывают эффекта на фенотип, поскольку изменение нуклеотидного состава не изменяет последовательность или количество белка либо, поскольку изменение аминокислотной последовательности не оказывает эффекта на функцию белка.
- Некоторые мутации влияют на функцию белка, однако практически не влияют на фенотип, поскольку белок сохраняет достаточную активность.

Мутации в генах могут оказывать различное влияние на фенотип (рис. 2.4).

Одним из способов определения функции гена является создание мутантного организма, дефицитного по данному гену. Мутация, полностью выключающая ген (чаще всего в результате удаления гена или его существенной части), называется **нуль-мутацией**. Если ген жизненно важен для организма, нуль-мутация будет летальной.

Для установления эффекта, оказываемого геном на фенотип, важно охарактеризовать нуль-мутанта. Если мутация не влияет на фенотип, это может быть просто **слабая мутация** (когда синтезируемый мутантный продукт достаточно активен для того, чтобы выполнять свою функцию, даже если его активность несколько ниже или несколько иная, чем у дикого типа). С другой стороны, если нуль-мутант не влияет на фенотип, можно с уверенностью утверждать, что кодируемая геном функция либо не важна, либо продублирована другим геном.

Нуль-мутации или мутации, влияющие на выполняемую геном функцию (но не обязательно полностью элиминирующие ее), называют мутациями с **потерей функции** (loss-of-function). Такие мутации рецессивны (как в примере на рис. 2.2). А иногда мутации приводят

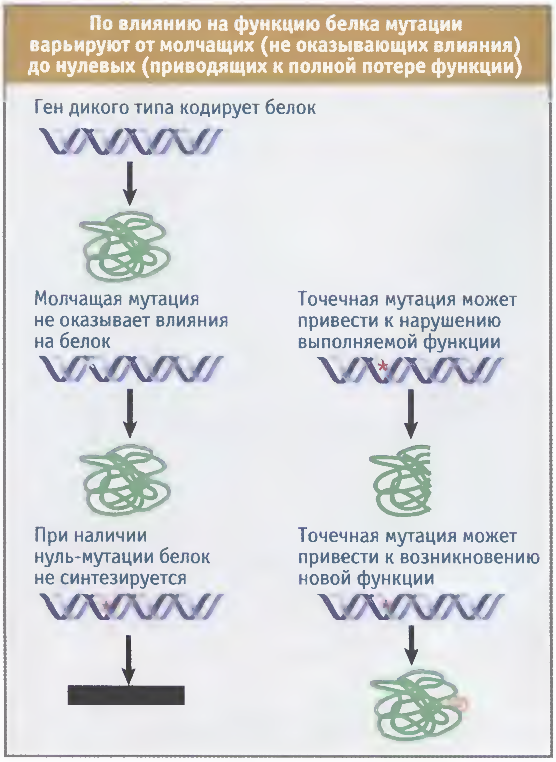


РИС. 2.4 Мутации, не изменяющие последовательность или функцию белка, называют молчащими. Мутации, приводящие к полной потере белком активности, называют нулевыми. Точечные мутации, приводящие к частичной потере функции, рецессивны, а усиливающие функцию — доминантны

к тому, что белок в результате изменения приобретает новую функцию; такие мутации называют мутациями с **приобретением функции** (gain-of-function). Такие мутации доминантны.

Не все мутации в гене приводят к детектируемым изменениям фенотипа. Мутации, не приводящие к очевидному эффекту, называют **молчащими.** Их, в свою очередь, делят на два типа: первые представляют собой нуклеотидные замены в последовательности ДНК, не приводящие к изменениям в аминокислотном составе соответствующего белка. Второй тип приводит к изменению аминокислотного состава, однако не приводит к изменению активности белка (такие мутации еще называют **нейтральными заменами**).

2.5 Локус может иметь различные мутантные аллели

Ключевые понятия:

- Существование множества аллелей определяет существование гетерозигот, содержащих любую комбинацию двух аллелей

Поскольку к рецессивным мутациям относят любые из них, негативно влияющие на синтез активного белка, именно такие мутации должны возникать наиболее часто. Многочисленные замены аминокислот могут изменить структуру белка достаточно сильно для того, чтобы нарушить его функцию.

Различные варианты одного и того же гена называют **множественными аллелями**. Существование множественных аллелей делает возможным появление гетерозиготных организмов, несущих два разных аллеля одного гена. При этом взаимодействие между аллелями может принимать различные формы.

В простейшем случае ген дикого типа кодирует нормально функционирующий белок, а мутантный аллель (аллели) кодирует нефункциональный белок.

Однако довольно часты случаи, когда множественные аллели приводят к формированию нескольких фенотипов. Например, дикий тип локуса *white* у *Drosophila melanogaster* необходим для развития нормального красного цвета глаз. Свое название локус получил благодаря эффекту, оказываемому нуль-мутацией — у гомозиготных по мутации мух глаза белые. Для разделения аллелей дикого типа и мутантных аллелей после названия локуса ставят надстрочный знак «+» (w^+ обозначает аллель дикого типа для дикого фенотипа глаз у *Drosophila melanogaster*). Иногда аллель дикого типа обозначают одним знаком «+», а мутантные аллели обозначают названием локуса.

Полностью дефектная форма гена (или мутантный фенотип) обозначают надстрочным знаком «–». Для обозначения множества мутантных аллелей, оказывающих различные эффекты, используют надстрочные знаки, такие как w^i или w^a.

Аллель w^+ в гетерозиготном состоянии является доминантным по отношению к любому другому аллелю. Для приведенного примера с окраской глаз у мух выявлено множество различных мутантных аллелей. На РИС. 2.5 приведен пример, демонстрирующий некоторые из них. Несмотря на то что отдельные аллели не влияют на цвет глаз, большинство из них обладают фенотипическим эффектом. Каждый из этих (мутантных) аллелей содержит изменение в гене, которое лишь частично подавляет его функцию, а остаточное влияние выражается в проявлении характерного фенотипа (большинство аллелей *w* оказывают влияние на количество пигмента в глазу). В данном случае аллели названы в соответствии с цветом глаз, формирующимся у гомозиготы (по данному аллелю). Приведенные на рисунке варианты перечислены в порядке уменьшения количества пигмента (однако некоторые варианты, как, например, w^{sp}, влияют не на количество, а на локализацию пигмента).

Когда в популяции присутствует множество аллелей данного гена, конкретный организм может быть гетерозиготой по двум разным аллелям. Фенотип такого организма зависит от природы остаточной активности каждого аллеля. Взаимовлияние двух му-

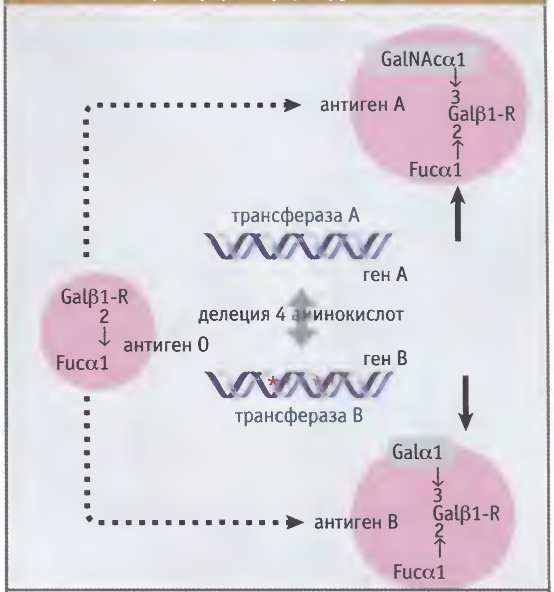
Каждому аллелю соответствует определенный фенотип

Аллель	Фенотип гомозиготы
w^{+}	красные глаза (дикий тип)
w^{bl}	глаза цвета крови
w^{ch}	вишневые
w^{bf}	темно-желтые
w^{h}	медовые
w^{a}	абрикосовые
w^{e}	цвета красителя эозина
w^{i}	цвета слоновой кости
w^{z}	лимонные
w^{sp}	пестрые, цвет варьирует
w^{1}	белые (цвет отсутствует)

РИС. 2.5 **Локус *w* представлен серией аллелей, определяющих спектр фенотипов от дикого типа до полного отсутствия пигмента**

тантных аллелей не отличается от взаимодействия мутантного аллеля и аллеля дикого типа: один из аллелей может быть доминантным, проявлять неполное доминирование или кодоминирование.

Галактозилтрансфераза формирует антигены А и В



Фенотип	Генотип	Активность
0	00	Нет
A	A0 или AA	N-Ac-Gal трансфераза
B	B0 или BB	Gal трансфераза
AB	AB	Gal-N-Ac-Gal трансфераза

РИС. 2.6 **Локус AB0 кодирует галактозилтрансферазу, специфичность которой определяет группу крови**

2.6 Локус может иметь более одного аллеля дикого типа

Ключевые понятия:

- Локус может быть представлен несколькими часто встречающимися аллелями, в этом случае среди них сложно выделить едниственный дикого типа

У конкретного генетического локуса может быть более одного аллеля дикого типа. Ярким примером такой ситуации служит наследование групп крови у человека. Угнетенная функция в данном случае обозначена нулевой группой крови, а взаимодействие аллелей показано на РИС. 2.6.

Антиген 0 (или Н) продуцируется у всех людей и представляет собой определенную углеводную группу, присоединяемую к белкам. Локус АВ0 кодирует фермент галактозилтрансферазу, присоединяющий сахар к антигену 0. Специфичность этого фермента и определяет группу крови. Аллель В продуцирует фермент, который использует кофактор УДФ-галактозы для создания антигена В. Аллель А кодирует фермент, использующий кофактор УДФ-N-ацетилгалактозу для производства антигена А. А- и В-аллели трансферазы отличаются четырьмя аминокислотами, которые, вероятно, и оказывают влияние на способность распознавать тип кофактора. Аллель 0 содержит небольшую делецию, которая элиминирует активность фермента и модификации антигена 0 не происходит.

Приведенная схема объясняет, почему аллели А и В доминантны у гетерозигот А0 и В0: соответствующая активность трансферазы создает А или В антиген. Аллели А и В проявляют у гетерозигот АВ так называемое кодоминирование, поскольку проявляют активность обе трансферазы. Гомозиготы 00 являются нулевыми, у них нет активного фермента, таким образом, они дефицитны по обоим антигенам.

Ни А-, ни В-аллель не могут считаться уникальными аллелями дикого типа, поскольку они проявляют активность, отличную от потери или приобретения функции. Случаи, когда в популяции присутствует множество функциональных аллелей, обозначают термином **полиморфизм** (см. разд. 4.3 «Геномы отличаются огромным разнообразием»).

2.7 Рекомбинация — это результат физических перестроек ДНК

Ключевые понятия:

- Рекомбинация — это результат процесса, называемого кроссинговером; он происходит с образованием специальных структур — хиазм и вовлекает две из четырех хроматид.
- Процесс рекомбинации включает разрыв и воссоединение цепей ДНК с формированием участков гибридной ДНК.

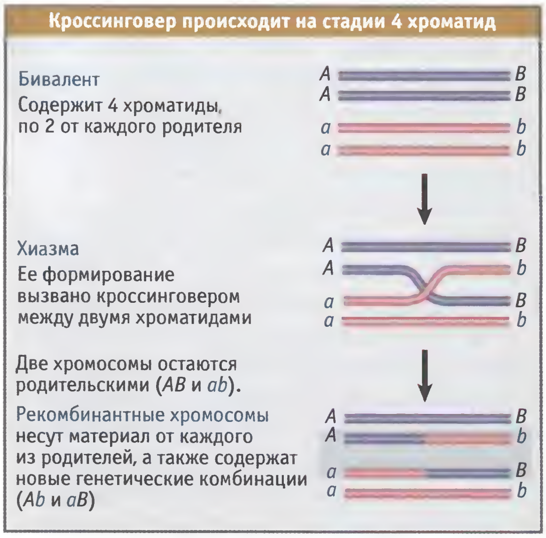


РИС. 2.7 Для возникновения рекомбинантов должна сформироваться хиазма

Генетическая рекомбинация приводит к образованию новой комбинации аллелей, что происходит в каждом поколении у диплоидных организмов. Как было сказано выше, каждая из двух хромосом может нести разные аллели в одном и том же локусе, и обмен частями хромосом (в процессе рекомбинации) приводит к образованию новых, рекомбинантных, хромосом, отличающихся от материнских.

Рекомбинация — это результат физических перестановок хромосомного материала. Этот процесс реализуется в ходе мейоза (деления клеток, в результате которого появляются гаплоидные гаметы) и называется кроссинговером. Как и в любое другое деление, в мейоз клетка вступает после удвоения хромосом и, таким образом, содержит четыре копии каждой хромосомы. На ранних стадиях мейоза все четыре копии связаны друг с другом, образуя структуру, называемую **бивалентом**. На данной стадии каждую из копий хромосомы называют **хроматидой**. Обмен генетическим материалом происходит между хроматидами, соединенными в бивалент.

В ходе кроссинговера образуется структура, называемая хиазмой. Ее схема представлена на РИС. 2.7. Хиазма — это место, в котором две хроматиды одного бивалента разрываются в гомологичных местах и соединяются крест-накрест, в результате чего образуются две новые хроматиды. Каждая новая хроматида с одной стороны разрыва содержит материал, полученный от первой хроматиды, а с другой стороны разрыва — от второй. Обе вновь образованные рекомбинантные хроматиды имеют так называемую реципрокную структуру (т. е. являющуюся результатом комбинации отцовского и материнского генетического материала). Описанное рекомбинационное событие принято называть **разрыв и воссоединение**. Из приведенного объяснения становится понятно, почему единичный акт рекомбинации приводит к появлению только двух (а не четырех) рекомбинантных хромосом: в процесс рекомбинации вовлечены только две из четырех соединенных в бивалент хроматид.

Процесс рекомбинации обусловливают общие принципы комплементарности цепей ДНК. Каждая из хроматид состоит из длинного дуплекса ДНК (рис. 2.7). В ходе рекомбинации необходимо очень точно распознавать места воссоединения рекомбинирующих молекул ДНК, в противном случае часть генетического материала будет утеряна. Механизм точного соответствия мест рекомбинации основан именно на принципах комплементарности.

В ходе рекомбинации две молекулы ДНК на определенном сравнительно коротком участке обмениваются своими цепями. В результате получается участок **гибридной ДНК** (РИС. 2.8), на котором цепь одного из ранее существовавших ДНК-дуплексов спарена с цепью другого дуплекса. Описываемый процесс включает и ряд других стадий (например, цепи должны быть каким-то образом разорваны и воссоединены), которые более подробно рассмотрены в гл. 20 «Системы репарации», однако принцип комплементарности цепей ДНК является основополагающим принципом гомологичной рекомбинации. Несмотря на то что на рисунке показаны только некоторые стадии процесса, видно, что участок гибридной ДНК образуется на промежуточной стадии рекомбинации — в момент, когда происходит перенос одноцепочечного участка ДНК с одного дуплекса на другой. Каждый из двух полученных в итоге рекомбинантов состоит из ДНК одной из исходных родительских хроматид (слева), соединенного с ДНК другой родительской хроматиды (справа). На рис. 2.7

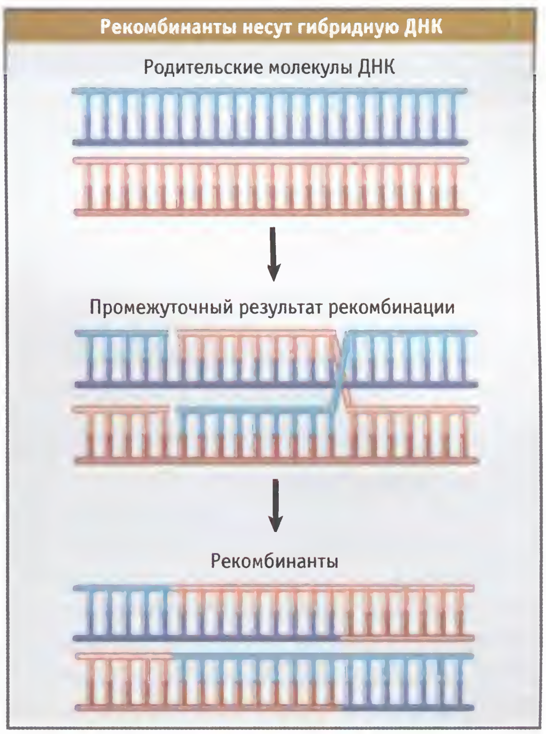


РИС. 2.8 Рекомбинация происходит в результате спаривания комплементарных цепей двух родительских дуплексов ДНК

каждый дуплекс ДНК соответствует одной из хроматид, вовлеченных в рекомбинацию.

Молекулы ДНК, принимающие участие в гомологичной рекомбинации, должны иметь достаточно схожие (комплементарные) последовательности. Настолько схожие, чтобы могла образоваться гибридная ДНК. Если две родительских хромосомы окажутся идентичны на участке обмена цепями, то сформируется идеальная гибридная ДНК. Однако рекомбинация может происходить даже в том случае, если между участками родительских ДНК есть незначительные отличия. В этом случае в гибридной ДНК возникнут неспаренные основания. Исправление таких несовпадений — это еще один аспект генетической рекомбинации (см. гл. 20 «Системы репарации»).

2.8 Генетический код триплетен

Ключевые понятия:

- Генетический код считывается триплетами нуклеотидов (кодонами).
- Триплеты не перекрываются и начинают считываться с определенной стартовой точки.
- Мутации, представляющие собой вставку или удаление отдельных нуклеотидов, приводят к сбою в считывании триплетов на участке ДНК после мутации.
- Сочетание мутаций, приводящих к добавлению или удалению сразу трех (или кратного трем числа) нуклеотидов, добавляет или убирает одну (несколько) аминокислот в белке, но не меняет рамку считывания триплетов после места расположения мутации.

Большинство генов несут информацию о строении белков, а каждый белок состоит из последовательности аминокислот. Эта концепция была разработана Сэнгером в 50-х годах двадцатого века на примере инсулина. Знание о том, что гены состоят из ДНК, позволяет нам понять, как последовательность нуклеотидов связана с последовательностью аминокислот в белке.

Особенностью ДНК является *слабая зависимость ее пространственной структуры от последовательности нуклеотидов*. Последовательность нуклеотидов в ДНК важна не потому, что она непосредственно формирует структуру белка, а потому, что она кодирует последовательность аминокислот, из которых состоит белок. Последовательность ДНК связана с кодируемым ею белком через **генетический код**.

Структура и/или ферментативная активность любого белка во многом зависит от его первичной структуры (последовательности составляющих его аминокислот). Ген, определяющий аминокислотную последовательность белка, несет всю необходимую информацию для создания функциональной полипептидной цепи. Таким образом, ген как структурная единица может быть реализован в виде бесчисленного количества полипептидов.

Все белки клетки ответственны за каталитические и структурные функции, которые в итоге формируют фенотип. Помимо собственно последовательностей, кодирующих белки, ДНК содержит участки, узнаваемые регуляторными молекулами (как правило, тоже белками). На данных участках функция ДНК также определяется ее последовательностью, но уже не в виде триплетного кода. Два упомянутых типа регионов — участки, непосредственно кодирующие белки, и регуляторные последовательности — в сумме составляют генетическую информацию.

Генетический код расшифровывается сложным аппаратом, интерпретирующим последовательность нуклеиновых кислот. Этот аппарат ответственен за распознавание значащих фрагментов ДНК. На каждом участке ДНК, как правило, только одна из двух цепей кодирует белок, поэтому генетический код записывают как последовательность оснований (а не пар оснований).

Генетический код считывается группами по три нуклеотида. Каждая такая группа (триплет) соответствует одной аминокислоте и называется **кодоном**. Ген состоит из серии кодонов, которые последовательно считываются, начиная со стартовой точки и заканчивая точкой терминации. Последовательность нуклеотидов, кодирующих белок, считывается в направлении от 5′- к 3′-концу кодирующей цепи ДНК и соответствует последовательности аминокислот в полипептиде в направлении от N-конца к C-концу.

Генетический код считывается с *неперекрывающихся триплетов, начиная с фиксированной стартовой точки*.

- Это означает, что каждый кодон состоит из трех нуклеотидов и каждый следующий кодон представлен новым триплетом нуклеотидов (триплеты не перекрываются).
- Сборка белка должна начинаться в *фиксированной стартовой точке*, кодирующей N-конец белка, и заканчиваться в точке, соответствующей C-концу, таким образом, чтобы различные части кодирующей последовательности не могли быть прочитаны независимо.

Знание структуры генетического кода позволяет предсказать эффект от двух типов мутаций. Если некоторый участок ДНК считывается последовательно и триплетами, например:

UUU	AAA	GGG	CCC	(кодоны)
ак1	ак2	ак3	ак4	(аминокислоты)

то точечная мутация приведет к изменению только одной аминокислоты. Например, замена A на любое другое основание (X) приведет к замене аминокислоты 2 на аминокислоту 5:

UUU	XAA	GGG	CCC
ак1	ак5	ак3	ак4

поскольку будет изменен только второй кодон.

Однако вставка или делеция одного основания приведет к изменению триплетов на протяжении всей сле-

дующей после мутации последовательности. Изменения такого рода называют **сдвигом рамки считывания**. Вставка (инсерция) может выглядеть следующим образом:

UUU XAA AGG GCC C
ак1 ак5 ак6 ак7

Так как новая последовательность триплетов отличается от старой, то и вся последовательность аминокислот в белке после точки мутации изменится. В результате белок может перестать выполнять свою функцию.

Сбой рамки вызывают **акридины** — соединения, которые связываются с ДНК и нарушают ее двухцепочечную структуру, приводя к вставке или делеции оснований в процессе репликации. Каждое мутационное событие, спровоцированное акридином, приводит к добавлению или удалению одной пары нуклеотидов из исходной последовательности ДНК.

В том случае, если мутация вызвана вставкой одного нуклеотида, возвращение к дикому типу может произойти в результате делеции этого нуклеотида. Однако такое возвращение может быть вызвано и делецией другого основания, находящегося вблизи от места расположения инсерции. Комбинации подобного рода мутаций наглядно подтверждают справедливость представлений о структуре генетического кода.

РИС. 2.9 иллюстрирует свойства мутаций, вызывающих сдвиг рамки. Инсерция или делеция вызывает изменения во всей последовательности аминокислот белка, закодированной участком после мутации. Однако комбинация инсерции и делеции приводит к неправильному считыванию кода только в пределах последовательности, расположенной между ними; правильное считывание возобновляется сразу после этого участка.

Проведенный в 1961 г. генетический анализ акридиновых мутаций в регионе *rII* фага T6 показал, что все мутации можно отнести к одной из двух групп: «+» или «–». Сам по себе каждый тип мутаций приводит к сбою рамки считывания: «+» путем добавления основания, а «–» путем делеции. Комбинации двойных мутантов каждого типа (++) и (––) также демонстрируют характерный для данной мутации эффект. Комбинации типа (+–) или (–+) подавляют друг друга, и в таких случаях одна из мутаций описывается как **супрессор сбоя рамки**, вызываемого другой мутацией. (В данном контексте термин «супрессор» используется в необычном для него значении, поскольку вторая мутация находится в том же гене, что и первая.)

В соответствии с приведенными результатами экспериментов генетический код должен считываться как последовательность с фиксированной стартовой точкой. Таким образом, инсерции или делеции могут компенсировать друг друга, в то время как двойные инсерции или двойные делеции остаются мутантными. Однако приведенные выше эксперименты не позволяют понять, из скольких нуклеотидов состоит кодон.

Оказалось, что дикий фенотип проявляется только у тройных мутантов (+++) и (–––), тогда как остальные комбинации числа вставок или делеций остаются мутантными. То, что инсерция или делеция в три основания ведет к добавлению или вырезанию одной аминокислоты доказывает, что код считывается триплетами. Измененная часть белка будет при этом ограничена участком между первым и третьим мутационными сайтами (рис. 2.9).

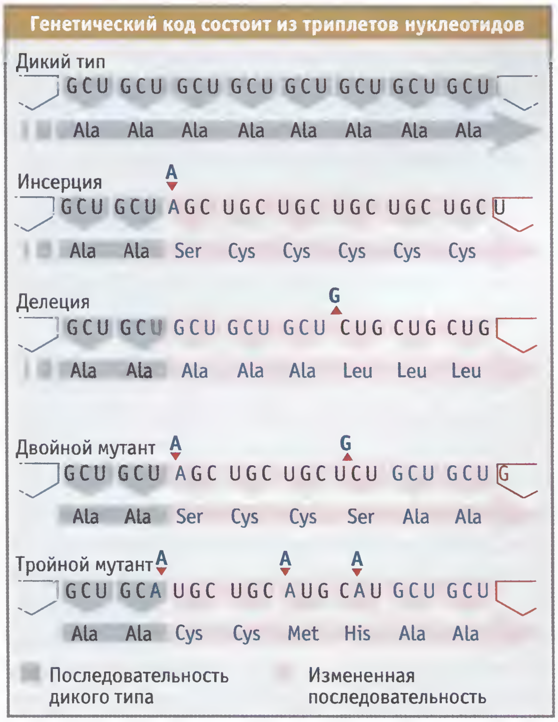


РИС. 2.9 Мутации, приводящие к сдвигу рамки считывания, подтверждают гипотезу о том, что генетический код считывается триплетами начиная с определенной стартовой точки

2.9 Существует три возможных рамки считывания каждой последовательности

Ключевые понятия:

- Как правило, транслируется только одна из трех рамок считывания, остальные заблокированы сигналами терминации трансляции.

Если основу генетического кода составляют неперекрывающиеся триплеты, то в зависимости от стартовой точки трансляции должно быть три возможных способа перевода любой нуклеотидной последовательности в белок, то есть три **рамки считывания**. Для последовательности

A C G A C G A C G A C G A C G A C G

возможны три рамки считывания:

```
   [ACG] ACG ACG ACG ACG ACG
A  [CGA]-CGA CGA CGA CGA CG
AC [GAC] GAC GAC GAC GAC G
```

Рамка считывания, состоящая кодирующих аминокислоты триплетов, называется **открытой рамкой считывания** (open reading frame, ORF). Транслируемая в белок последовательность включает рамку считывания, состоящую из начального **старт-кодона** (AUG), серии кодирующих аминокислоты триплетов и одного из трех **стоп-кодонов** (см. гл. 7 «Матричная РНК»).

Если внутри рамки считывания встречаются стоп-кодоны, то белок с нее не считывается. В таких случаях говорят, что рамка считывания **блокирована**. Последовательность нуклеотидов не может выполнять белок-кодирующую функцию, если она блокирована во всех трех рамках считывания.

Чтобы понять, существует ли на данном участке ДНК открытая рамка считывания, необходимо проанализировать все возможные рамки считывания. Обычно на одном участке ДНК открытой является только одна из трех возможных рамок считывания. На РИС. 2.10 приведен пример последовательности, для которой существует лишь одна продолжительная рамка, а альтернативные рамки блокированы частыми стоп-кодонами.

Вероятность возникновения не кодирующей белок (случайной), сравнительно длинной открытой рамки считывания тем меньше, чем она длиннее, а естественный отбор против накапливания в некодирующей рамке стоп-кодонов не происходит. Можно утверждать, что если мы обнаружили сравнительно длинную открытую рамку считывания без стоп-кодонов, то данная рамка считывания кодирует некий белок. Открытую рамку считывания, для которой не обнаружен продукт, иногда называют неидентифицированной рамкой считывания (unidentified reading frame, URF).

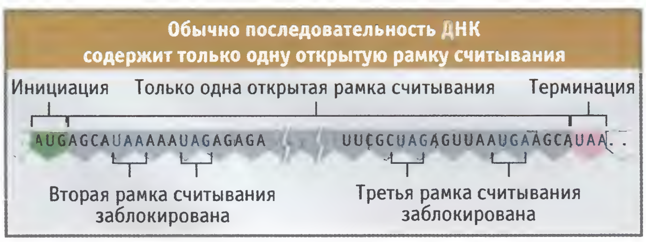


РИС. 2.10 Открытая рамка считывания начинается с кодона AUG и продолжается триплетами вплоть до стоп-кодона. Заблокированные рамки считывания прерваны стоп-кодонами

2.10 Прокариотические гены колинеарны их белкам

Ключевые понятия:

- Прокариотический ген включает участок из $3N$ нуклеотидов, кодирующих N аминокислот.
- Ген, мРНК и белок колинеарны друг другу.

Сравнивая нуклеотидную последовательность гена с последовательностью аминокислот в белке, можно определить, **колинеарны** ли ген и белок (то есть существует ли линейное соответствие между последовательностью нуклеотидов в гене и последовательностью аминокислот в белке). Например, такое линейное соответствие доказано для бактерий и их вирусов. Каждый бактериальный или вирусный ген состоит из протяженного участка ДНК, длина которого напрямую связана с числом кодируемых им аминокислот. В соответствии с устройством генетического кода ген, состоящий из $3N$ пн, должен кодировать белок, состоящий из N аминокислот.

Пропорциональность бактериальных генов и их продуктов означает, что физическая карта ДНК в точности совпадет с аминокислотной картой белка. Но насколько эти карты совпадают с картой рекомбинации?

Первые исследования колинеарности генов и белков проводили на примере гена триптофансинтетазы *E. coli*. Генетические расстояния измеряли в единицах рекомбинации между мутациями, а расстояния в белке — по числу аминокислот, разделяющих сайты мутаций. На РИС. 2.11 приведено сравнение двух рекомбинационных карт. Порядок расположения семи сайтов мутаций совпадал с порядком соответствующих аминокислотных замен, а рекомбинационные расстояния были очень похожи на расстояния между аминокислотными заменами в белке. Карта рекомбинации несколько преувеличивает расстояния между некоторыми мутациями, поскольку она искажена относительно физической карты.

Карта рекомбинации указывает еще на два важных момента, касающихся организации гена. Мутация может приводить к замене аминокислоты дикого типа на другую аминокислоту. Если две таких мутации никогда не рекомбинируют, то они должны представлять собой различные точечные мутации, расположенные в одном и том же месте на ДНК. Если на генетической карте две мутации разделяются, но при этом затрагивают одну аминокислоту (сходящиеся линии на карте), очевидно, они должны представлять собой точечные мутации в одном триплете. Следовательно, единица генетической рекомбинации (в действительности — 1 пн) меньше, чем единица, кодирующая аминокислоту (3 пн).

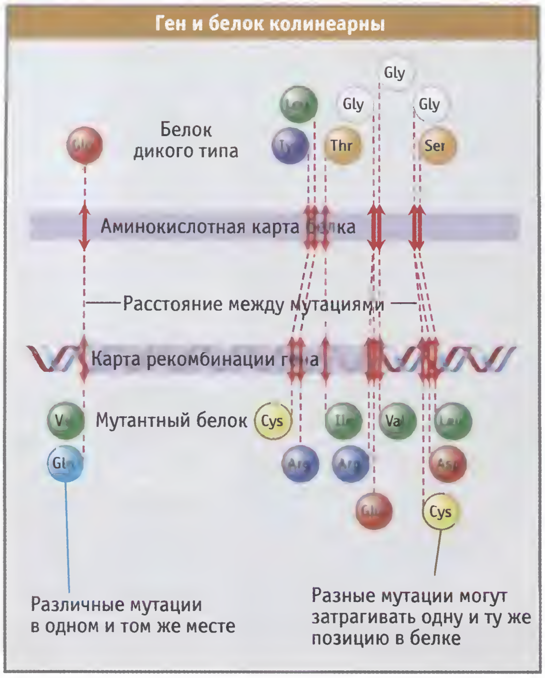


РИС. 2.11 Карта рекомбинации гена триптофансинтетазы соотносится с аминокислотной последовательностью белка

2.11 Синтезу белка предшествует ряд процессов

Ключевые понятия:

- Прокариотический ген экспрессируется через транскрипцию (синтез мРНК) и трансляцию (мРНК) с образованием белка.
- У эукариот ген может содержать внутренние регионы, не представленные в белке.
- Внутренние регионы удаляются из эукариотического РНК-транскрипта посредством РНК-сплайсинга. В итоге эукариотическая мРНК также оказывается колинеарна белку.
- Зрелая молекула мРНК содержит кодирующий регион и нетранслируемые 5′-регион и 3′-регион.

Если сопоставить последовательности гена и белка, окажется, что белок-кодирующая последовательность (расположенная между точками, соответствующими концам полипептидной цепи) является только частью гена. Поскольку ген не напрямую транслируется в белок, а посредством **матричной (информационной) РНК**, сокращенно мРНК, промежуточной в синтезе белка нуклеиновой кислоты (подробнее см. гл. 7 «Матричная РНК»), существуют участки гена, не кодирующие полипептидную цепь, но входящие в структуру мРНК.

Процесс синтеза мРНК происходит по уже знакомому нам механизму комплементарного спаривания оснований (так же, как и при репликации ДНК), с тем лишь отличием, что синтез мРНК идет лишь по одной из цепей ДНК. На РИС. 2.12 показано, что последовательность мРНК комплементарна одной из цепей ДНК и идентична второй из цепей (за исключением замены T на U). Видно, что последовательность верхней цепи ДНК (направленая от 5′-конца к 3′-концу) соответствует последовательности РНК.

Процесс, в результате которого ген дает начало белку, называют экспрессией гена. У бактерий экспрессия генов осуществляется в две стадии. Во время первой стадии, называемой **транскрипцией,** с одной из цепей ДНК синтезируется копия мРНК. Во время второй стадии мРНК **транслируется** в белок. Таким образом, трансляция — это процесс, в ходе которого последовательность мРНК, состоящая из триплетов, кодирующих аминокислоты, формирует соответствующий белок. Часть мРНК, называемая **кодирующим регионом**, состоит из последовательности нуклеотидов, кодирующих последовательность аминокислот в белке. Кроме кодирующего региона мРНК, как правило, содержит дополнительные регионы по концам молекулы; эти последовательности не принимают непосредственного участия в кодировании белка. 5′-нетранслируемый регион в англоязычной литературе называют **лидером**, а 3′-нетранслируемый регион — **трейлером**.

Ген включает в себя полную последовательность мРНК. Иногда в некодирующих регионах обнаруживают мутации, нарушающие функцию гена, из чего можно сделать вывод, что эти регионы составляют важную часть гена.

На РИС. 2.13 продемонстрирована ситуация, в которой ген, представленный протяженным участком ДНК, продуцирует определенный белок. Помимо белок-кодирующей последовательности он несет по концам дополнительные некодирующие регионы.

Поскольку цитоплазма у бактерий не отделена от геномной ДНК (у бактерий отсутствует ядерная оболочка), транскрипция и трансляция происходят у них в одном и том же компартменте клетки (РИС. 2.14).

В отличие от бактерий у эукариот транскрипция осуществляется в ядре. Для того чтобы транслировать

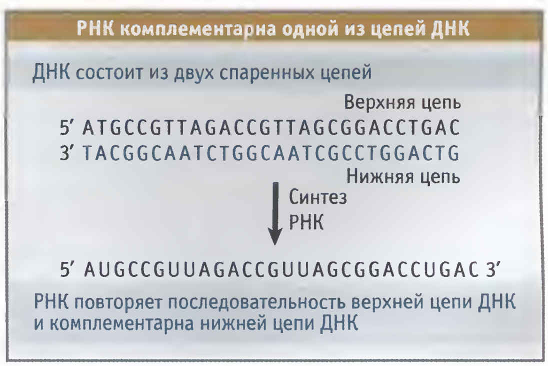


РИС. 2.12 В ходе синтеза РНК одна из молекул ДНК используется в качестве матрицы для комплементарного спаривания нуклеотидов

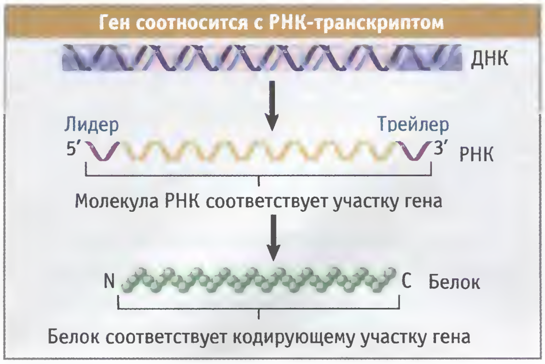


РИС. 2.13 По длине ген может превышать кодирующую белок последовательность

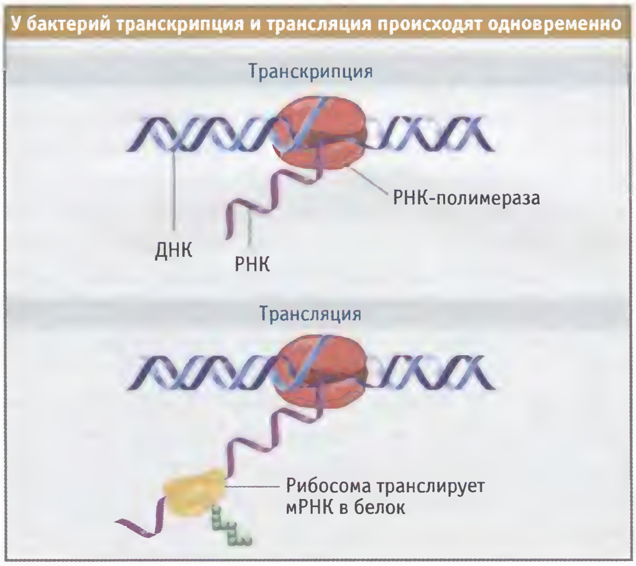


РИС. 2.14 У бактерий транскрипция и трансляция происходят в одном компартменте клетки

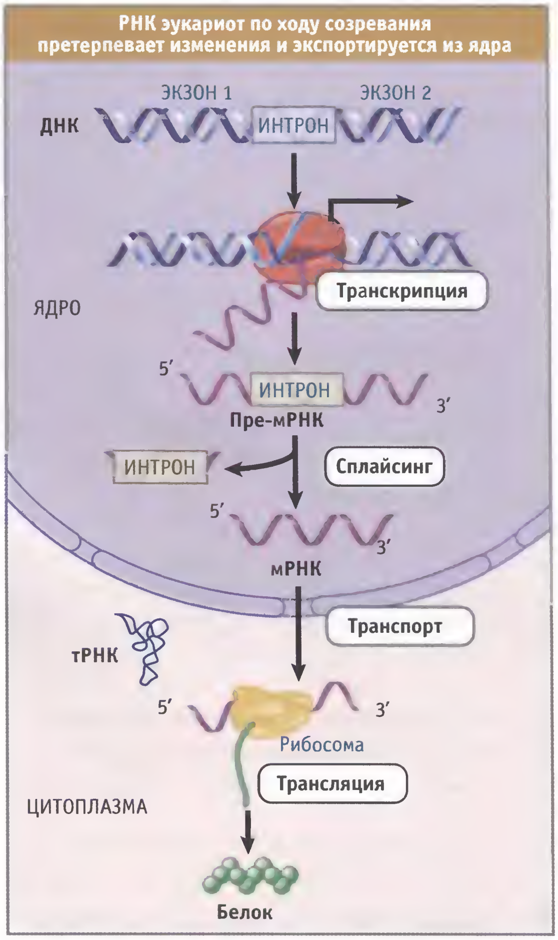


РИС. 2.15 Процесс экспрессии гена включает несколько стадий

мРНК в белок, необходимо *транспортировать* РНК в цитоплазму. Многие гены эукариот (и большинство генов высших эукариот) содержат внутренние регионы, не кодирующие белок — интроны. В (безынтронных) генах эукариот РНК-транскрипт по сути является мРНК (как у бактерий). В интронсодержащих генах синтезируемая на ДНК молекула представляет собой более длинный предшественник мРНК (**пре-мРНК**), и чтобы получить зрелую мРНК, из нее нужно удалить интроны (сплайсировать).

Модификации РНК, происходящие после ее синтеза на ДНК, называют **процессингом РНК**. Самой важной стадией процессинга РНК является **сплайсинг**. Во время сплайсинга эти участки удаляются из пре-мРНК, формируя молекулу мРНК, содержащую протяженную открытую рамку считывания (см. рис. 3.1). Помимо сплайсинга в ходе процессинга происходят и другие важные события: модификация 5′- и 3′-концов пре-мРНК (рис. 7.16). Основные стадии экспрессии генов эукариот приведены на РИС. 2.15. Из рисунка видно, что процессы транскрипции и трансляции пространственно разделены: первый осуществляется в ядре, а второй — в цитоплазме.

Трансляцию осуществляет сложный ферментативный аппарат, включающий как белки, так и РНК. Ключевым участником процесса сборки белка является *рибосома* — большой макромолекулярный комплекс, состоящий из специальных РНК (*рибосомальные РНК*, обозначают как *рРНК*) и множества сравнительно небольших белков. В процесс распознавания аминокислот, соответствующих триплетам нуклеотидов, вовлечены так называемые *транспортные РНК* (*тРНК*); каждой аминокислоте соответствует как минимум один вид тРНК. Помимо этого, в процесс вовлечено множество вспомогательных белков. Более подробное описание процесса трансляции приводится в гл. 7 «Матричная РНК», сейчас

же достаточно понять, что рибосомы — это крупные РНК-белковые структуры, которые продвигаются вдоль мРНК (рис. 2.14).

Особо отметим, что РНК принимают участие в экспрессии генов не только как информационная матрица, но и в виде компонентов аппарата трансляции. рРНК и тРНК, так же, как и мРНК, кодируются генами и синтезируются в процессе транскрипции, однако последовательность их нуклеотидов не транслируется в белок.



2.16 В ДНК расположены регуляторные участки, с которыми могут связываться белки; экспрессия кодирующих участков осуществляется посредством синтеза РНК

2.12 Регуляторные белки — *транс*-действующие, а регуляторные участки ДНК — *цис*-действующие

Ключевые понятия:

- РНК и белки являются *транс*-действующими агентами, то есть могут воздействовать на любую из копий регулируемого ими гена в данной клетке.
- *Цис*-действующие мутации позволяют определять последовательности ДНК, являющиеся мишенями для *транс*-действующих РНК или белков. В отличие от РНК или белков *цис*-действующие мутации не перемещаются в пространстве, и могут оказывать влияние только на данную молекулу ДНК.

Важным шагом в формировании представлений о гене стало понимание того, что все его части должны располагаться на одном непрерывном участке (сайте) ДНК. В классической генетике принято говорить, что участки, расположенные на одной и той же молекуле ДНК, находятся в ***цис***-положении, а участки, расположенные на двух разных молекулах ДНК, — в ***транс***-положении. Таким образом, две мутации могут находиться в *цис-* (на одной молекуле) или в *транс-* (на разных молекулах) положениях. Чтобы определить, в каком положении в гене расположены две мутации, используют комплементационный тест (рис. 2.3). Теперь мы можем рассматривать различные вопросы (начиная с определения кодирующего региона гена и заканчивая описанием взаимодействия между его регуляторными элементами), используя терминологию *цис-* и *транс*-расположения.

Допустим, экспрессию некоторого гена контролирует белок, связывающийся с ДНК в районе этого гена. В примере, приведенном на РИС. 2.16, синтез РНК запускается только при условии связывания специального белка с определенным участком ДНК. Теперь предположим, что на участке связывания белка с ДНК возникнет мутация, в результате которой белок больше не узнает свою мишень. Следовательно, мутация приведет к отсутствию экспрессии данного гена.

Таким образом, ген может быть инактивирован как мутацией в кодирующем регионе, так и мутацией в регуляторном участке. Методами классической генетики такие мутации различить невозможно, поскольку обе они влияют на функционирование одного и того же аллеля (и, как следствие, обнаруживают одинаковые свойства при проведении комплементационного теста). Мутация в регуляторном регионе определяется комплементационным тестом так же, как и мутация в кодирующем регионе. На РИС. 2.17 показано, что дефект в

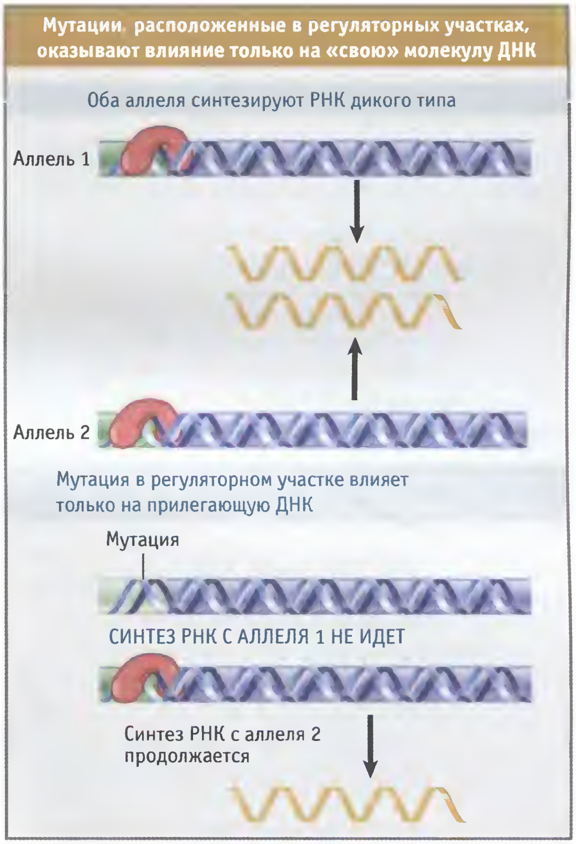


РИС. 2.17 Регуляторный участок (*цис*-действующий) влияет только на примыкающую к нему ДНК и не оказывает влияния на другой аллель

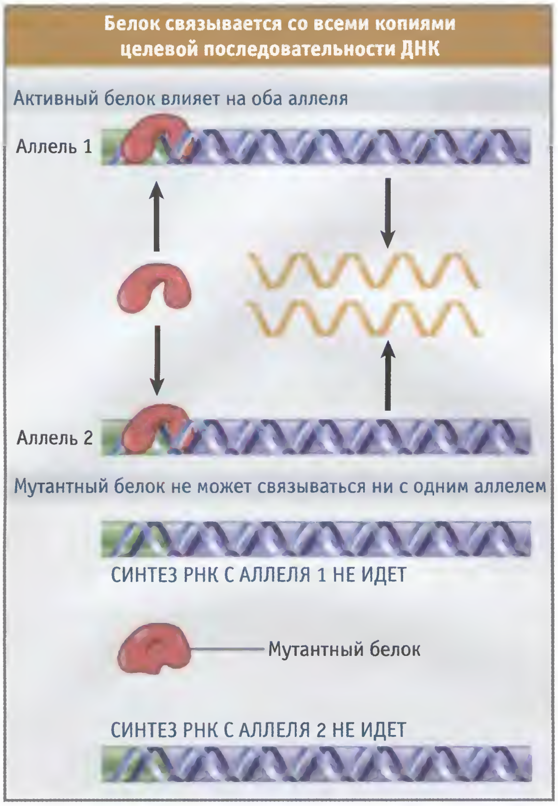


РИС. 2.18 Мутация белка влияет на оба аллеля гена, контролируемого этим белком (*транс*-действие)

регуляторном регионе *затрагивает только тот кодирующий регион, к которому он относится, и не влияет на экспрессию других аллелей.* Мутацию, влияющую только на свойства связанного с ней региона ДНК, называют *цис*-действующей.

В противовес поведению *цис*-действующих мутаций (рис. 2.17), можно привести пример с мутацией в гене, кодирующем собственно регуляторный белок. На **РИС. 2.18** показано, что отсутствие регуляторного белка (в результате мутации его гена) заблокирует экспрессию *обоих* присутствующих в клетке аллелей гена. Мутации такого рода называют *транс*-действующими.

Если посмотреть на приведенный пример с молекулярной точки зрения, мы увидим, что эффект от *транс*-действующей мутации должен выражаться через некоторого посредника (чаще всего — белковой природы), действующего на множественные мишени в клетке, в то время как *цис*-действующая мутация, оказывая эффект только на смежную с ней ДНК, *не экспрессируется в виде РНК или белка.*

2.13 Резюме

Хромосома состоит из непрерывного ДНК-дуплекса, содержащего множество генов. Каждый ген (цистрон) транскрибируется в РНК, которая, в свою очередь, транслируется в полипептидную цепь (если, конечно, этот ген кодирует белок). Принято говорить, что продукты гена (в виде РНК или белка) являются *транс*-действующими. Определить ген как некоторый регион ДНК можно с помощью комплементационного теста. Участок ДНК, регулирующий активность прилегающего к нему гена, является *цис*-действующим.

Соответствие последовательности гена (кодирующего белок) и последовательности аминокислот в белке определяются генетическим кодом.

Только одна из двух цепей ДНК кодирует белок. Одну аминокислоту кодирует последовательность из трех нуклеотидов, называемая кодоном. Кодирующая последовательность ДНК состоит из серии кодонов, считываемых начиная с фиксированной стартовой точки. Синтез белка происходит в одной из трех возможных рамок считывания.

У гена может быть много аллелей (вариантов) в популяции организмов данного вида. Рецессивные аллели, как правило, связаны с потерей функции кодируемого белка. Нуль-аллель соответствует полной потере функции жизненно важного белка. Доминантные аллели, как правило, связаны с приобретением полезных для организма функций.

Литература

2.8 Генетический код триплетен

Обзор

Roth, J. R. (1974). Frameshift mutations. *Annu. Rev. Genet.* 8, 319–346.

Исследование

Benzer, S. and Champe, S. P. (1961). Ambivalent rII mutants of phage T4. *Proc. Natl. Acad. Sci. USA* 47, 403–416.

Crick, F. H. C., Barnett, L., Brenner, S., and Watts-Tobin, R. J. (1961). General nature of the genetic code for proteins. *Nature* 192, 1227–1232.

2.10 Прокариотические гены колинеарны их белкам

Исследование

Yanofsky, C., Drapeau, G. R., Guest, J. R., and Carlton, B. C. (1967). The complete amino acid sequence of the tryptophan synthetase A protein (m subunit) and its colinear relationship with the genetic map of the A gene. *Proc. Natl. Acad. Sci. USA* 57, 2966–2968.

Yanofsky, C. et al. (1964). On the colinearity of gene structure and protein structure. *Proc. Natl. Acad. Sci. USA* 51, 266–272.

Прерывистый ген

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

3.1 Введение

3.2 Прерывистый ген состоит из экзонов и интронов

- Интроны удаляются в результате сплайсинга РНК, осуществляемого в *цис*-положении.
- Повлиять на последовательность белка могут только мутации, находящиеся в экзонах; тем не менее мутации в интронах также могут оказывать влияние на синтез белка путем воздействия на механизм процессинга РНК.

3.3 Эндонуклеазы рестрикции — это один из основных инструментов картирования ДНК

- Эндонуклеазы рестрикции можно использовать для расщепления ДНК на фрагменты.
- Перекрывающиеся участки фрагментов ДНК, полученных с помощью различных рестриктаз, позволяют создавать генетические карты.

3.4 Организация прерывистых генов консервативна

- С помощью рестрикционных карт или электронной микроскопии можно провести сравнение генов и их РНК-продуктов: отсутствующие в зрелых мРНК регионы генов и будут являться интронами.
- Сравнение гомологичных генов разных организмов позволяет сделать вывод о консервативности расположения интронов, однако длины аналогичных (по положению в гене) интронов могут значительно варьировать.
- Обычно интроны не кодируют белки.

3.5 Последовательности экзонов консервативны, а интронов — вариабельны

- При межвидовом сравнении родственных генов видно, что последовательности экзонов консервативны, тогда как интроны не похожи.
- Интроны эволюционируют быстрее, чем экзоны Отбор воздействует на интроны в значительно меньшей степени, поскольку большая часть их последовательности не влияет на функциональность белка.

3.6 Длина генов сильно варьирует

- У дрожжей многие гены непрерывны (не содержат интронов), тогда как у высших эукариот большинство генов имеют прерывистое строение.
- Обычно экзоны короче интронов и кодируют не более 100 аминокислот.
- У низших эукариот интроны сравнительно короткие, тогда как у высших эукариот их длина может достигать десятков тысяч пар нуклеотидов.
- Длина эукариотического гена в основном определяется длиной его интронов.

3.7 Некоторые последовательности ДНК кодируют более одного белка

- Возможность использования дополнительных кодонов, инициирующих и терминирующих транскрипцию, позволяет получить с одного гена два белка, один из которых соответствует части другого.
- Одна и та же последовательность ДНК может отвечать за синтез негомолгичных белков при условии, что информация считывается в разных рамках для двух перекрывающихся генов.
- Гомологичные белки, различающиеся наличием или отсутствием определенных регионов, могут возникать в результате альтернативного сплайсинга. Это может быть включение или исключение из зрелой РНК отдельных экзонов (или выбора между двумя альтернативными экзонами).

3.8 Как эволюционировали прерывистые гены?

- Основной вопрос эволюции: «Как возникли гены?» Имели они прерывистую структуру с самого начала или возникли как непрерывные последовательности и затем были «обогащены» интронами?
- Вероятно, подавляющая часть генов, кодирующих белки, возникла в виде прерывистой формы, однако гены, кодирующие рибосомальные и транспортные РНК, изначально могли быть непрерывны.
- Существует особый класс мобильных интронов, способных самостоятельно встраиваться в гены.

3.9 Некоторые экзоны напоминают последовательности, кодирующие целые белки

- В пользу того, что экзоны были строительным материалом эволюции и что первые гены имели прерывистую структуру, говорят следующие факты:
  - у довольно сильно отстоящих друг от друга видов гены имеют схожую экзон-интронную структуру;
  - существует большое число экзонов, последовательности которых могут быть сопоставлены с последовательностями, кодирующими целые белки;
  - существует гомология между экзонами, находящимися в разных генах.

3.10 Гены, относящиеся к одному семейству, имеют схожую организацию

- Схожесть черт у генов, относящихся к одной группе, позволяет сделать предположение о функции, выполняемой ими до разделения в эволюции.

- Все гены глобинов схожи по своей структуре. Они содержат три экзона и два интрона, что дает основания предполагать, что они произошли от одного предкового гена.

- В последнее время определение термина «ген» существенно изменилось. Утверждение «один ген — один белок» превратилось в «один белок — один ген».
- Позиционная информация также важна при развитии организма.

3.1 Введение

Простейшей формой гена является участок ДНК, колинеарный белковому продукту. Такую структуру имеют практически все бактериальные гены. В генах, относящихся к этому типу, протяженная последовательность, состоящая из 3N оснований, кодирует N аминокислот. Однако у эукариот в пределах кодирующего региона могут находиться дополнительные последовательности, прерывающие участок, кодирующий тот или иной белок. В ходе экспрессии гена эти последовательности удаляются из РНК. В результате получается мРНК, последовательность которой колинеарна последовательности белкового продукта в точном соответствии с правилами генетического кода.

Регионы ДНК, составляющие прерванный ген, делятся на две категории (РИС. 3.1):

- **Экзоны** — это участки ДНК, из которых состоит зрелая мРНК. По определению, ген начинается и заканчивается экзонами, соответствующими 5′- и 3′-концам РНК.
- **Интроны** — это внутренние регионы гена, удаляемые из первичного транскрипта и отсутствующие в зрелой РНК.

Порядок расположения экзонов одинаков для гена и кодируемой им РНК, но прерывистый ген длиннее, чем зрелая мРНК, поскольку такой ген содержит интроны.

Экспрессия прерывистых генов (в отличие от безинтронных генов) требует наличия дополнительного этапа. Сначала с ДНК прерывистого гена синтезируется копия РНК (первичный транскрипт), в точности соответствующая геномной последовательности. Эта РНК не может быть использована для синтеза белка — сначала из первичного транскрипта должны быть удалены интроны, формируя зрелую мРНК, состоящую только из экзонов. Этот процесс называется сплайсингом РНК (см. разд. 2.11 «Синтезу белка предшествует ряд процессов») и заключается в точном удалении интронов из первичного транскрипта с последующим ковалентным соединением концов РНК с формированием целой молекулы (см. гл. 26 «Сплайсинг и процессинг РНК»).

Ген представляет собой участок генома, ограниченный точками, соответствующими 5′- и 3′-концам зрелой мРНК. Мы знаем, что транскрипция начинается на 5′-конце мРНК, а заканчивается за пределами 3′-конца зрелой мРНК (поскольку 3′-конец образуется с результате разрезания пре-мРНК, см. разд. 26.19 «3′-концы мРНК формируются путем разщепления и полиаденилирования первичного транскрипта»). По обоим концам ген содержит регуляторные регионы, необходимые для инициации и (иногда) терминации его экспрессии.

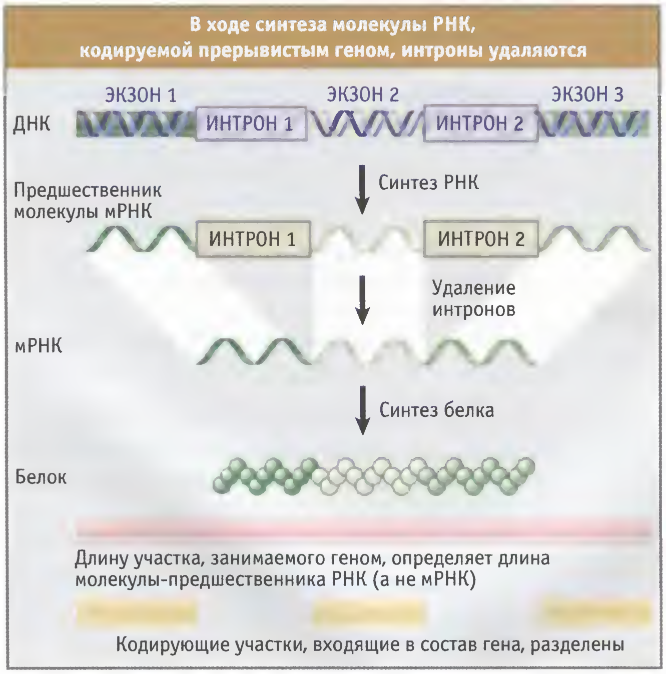


РИС. 3.1 Экспрессия прерывистых генов осуществляется посредством молекулы-предшественника РНК. В процессе объединения экзонов интроны удаляются. мРНК содержит только последовательности экзонов

3.2 Прерывистый ген состоит из экзонов и интронов

Ключевые понятия:

- Интроны удаляются в результате сплайсинга РНК, осуществляемого в *цис*-положении.
- Повлиять на последовательность белка могут только мутации, находящиеся в экзонах; тем не менее мутации в интронах также могут оказывать влияние на синтез белка путем воздействия на механизм процессинга РНК.

Как существование интронов изменяет наши представления о гене? После сплайсинга экзоны соединяются в том же порядке, в котором они представлены в ДНК. Таким образом, колинеарность гена и белка у эукариот определяется отдельными экзонами и соответствующими им участками полипептидной цепи. На РИС. 3.2 показано, что порядок



РИС. 3.2 В молекуле мРНК экзоны сохраняют тот же порядок, что и в молекуле ДНК, однако длина участков между двумя точками в экзонах, расположенных в гене, не совпадает с длиной участков между двумя соответствующими точками в мРНК или белке. В гене расстояние A–B меньше расстояния B–C, а в РНК и белке наоборот

расположения мутаций в гене соответствует порядку расположения аминокислотных замен в белке. Как видно из карты рекомбинации, генетические расстояния не имеют отношения к расстояниям между соответствующими точками в белке. Длина гена определяется не длинной мРНК, а длиной начальной (предшествующей) молекулы РНК.

Все экзоны гена расположены на одной молекуле РНК и их сплайсинг проходит как *внутри*молекулярная реакция. Экзоны, входящие в состав *разных* молекул РНК, обычно не соединяются. Таким образом, механизм сплайсинга исключает объединение последовательностей, представляющих разные аллели. Мутации, расположенные в разных экзонах одного гена, не могут дополнять друг друга и являются членами одной комплементационной группы.

Мутации, напрямую влияющие на последовательность белка, должны находиться в экзонах. Какой же эффект оказывают мутации в интронах? Интроны не являются частью зрелой мРНК. Следовательно, мутации в интронах не могут напрямую влиять на структуру белка. Однако они могут помешать производству зрелой молекулы мРНК (например, путем нарушения механизма сплайсинга данной РНК). Мутации подобного рода действуют только на тот аллель, в котором они расположены. Они не могут дополнять любые другие мутации данного аллеля и всегда являются частью той же комплементационной группы, что и экзоны.

Подавляющее большинство влияющих на сплайсинг мутаций являются однонуклеотидными заменами в местах соединения интронов и экзонов. Они могут вести к выпадению части экзона из продукта (в результате возникновения дополнительного интрона) или же спровоцировать сплайсинг в неправильном месте. Наиболее часто встречающееся следствие неправильного сплайсинга — это возникновение стоп-кодона и, как следствие, нарушение последовательности белка. Около 15% точечных мутаций, вызывающих заболевания у людей, нарушают процесс сплайсинга.

Не все эукариотические гены имеют прерывистую структуру. Некоторые из них, как и гены прокариот, в точности соответствуют своему белковому продукту. У дрожжей подавляющее большинство генов непрерывны (безынтронны). У высших эукариот большинство генов содержат интроны и обычно значительно длиннее суммы длин экзонов.

3.3 Эндонуклеазы рестрикции — один из основных инструментов картирования ДНК

Ключевые понятия:

- Эндонуклеазы рестрикции можно использовать для расщепления ДНК на фрагменты.
- Перекрывающиеся участки фрагментов ДНК, полученных с помощью различных рестриктаз, позволяют создавать генетические карты.

Характеристика структуры эукариотических генов стала возможной в результате развития методов физического картирования ДНК. Эти методы распространяются и на РНК, поскольку существуют сравнительно простые способы создания двухцепочечной ДНК-копии молекулы РНК. Физическая карта любой молекулы ДНК может быть получена путем расщепления молекулы в определенных местах и последующего определения размера получаемых фрагментов. Такое расщепление проводят с использованием эндонуклеаз рестрикции — ферментов, способных распознавать довольно короткие последовательности двухцепочечной ДНК в качестве мишени для разрезания.

Каждая рестриктаза нацелена на определенную мишень в двухцепочечной ДНК. Обычно это последовательность, состоящая из 4–6 пар оснований (сайт). Фермент разрезает ДНК во всех местах (сайтах), соответствующих специфичной для данной рестриктазы последовательности. Разные рестриктазы распознают разные последовательности. На сегодняшний день спектр рестриктаз (получаемых из огромного разнообразия бактерий) представлен чрезвычайно широко.

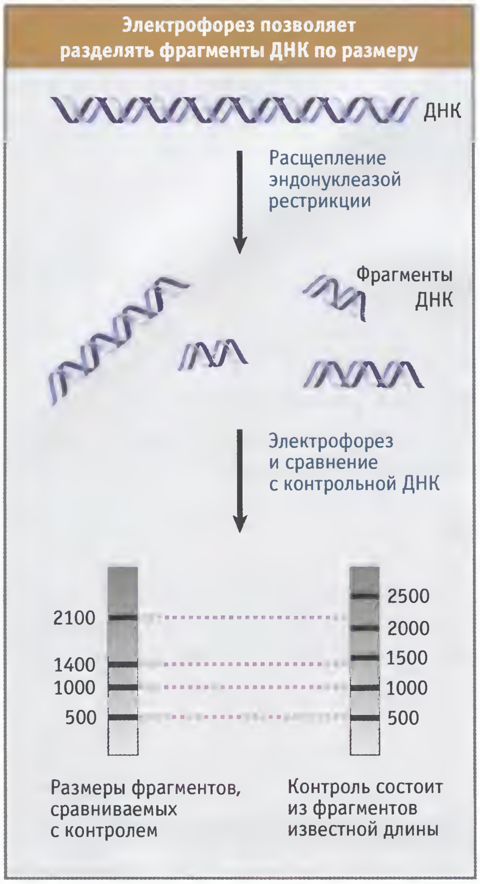


РИС. 3.3 Фрагменты, получившиеся в результате расщепления ДНК эндонуклеазами рестрикции, могут быть разделены в соответствии с их размером

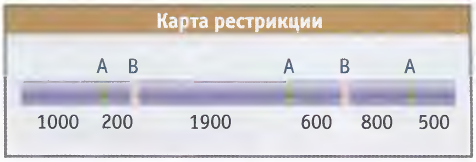


РИС. 3.4 Карта рестрикции представляет собой линейную последовательность участков ДНК, разделенных промежутками известной длины. С помощью отдельных фрагментов, сформированных в результате одиночных и двойных расщеплений, на карте можно расположить сайты ферментов А и В

Карта рестрикции представляет собой линейную последовательность участков, на стыках которых данная рестриктаза обнаружила свою мишень. На коротких дистанциях расстояние оценивают в **парах нуклеотидов** (пн). Более длинные участки ДНК и РНК измеряют в **тысячах пар нуклеотидов** (тпн). На уровне хромосом расстояния измеряют миллионами пар нуклеотидов, или **мега-парами** (1 **Мпн** = 10^6 пн).

Рестриктаза расщепляет молекулу ДНК на определенные фрагменты. Эти фрагменты можно разделить по размеру методом гель-электрофореза (РИС. 3.3). Расщепленную ДНК помещают в гель, приготовленный из агарозы или полиакриламида. При пропускании тока через гель фрагменты ДНК продвигаются от анода к катоду со скоростью, обратно пропорциональной молекулярному весу фрагмента. По ходу движения в геле фрагменты формируют серии полос, каждая из которых соответствует фрагменту определенного размера. Чем ближе к катоду расположена полоса в геле, тем меньше длина фрагмента.

Анализируя рестрикционные фрагменты ДНК, можно создать карту исходной молекулы, как показано на РИС. 3.4. Карта показывает места, в которых данная рестриктаза разрезала ДНК. Расстояния между сайтами рестрикции измеряют в парах нуклеотидов. *Таким образом, ДНК разделяется на серии фрагментов определенной длины, разделенных сайтами распознавания рестриктаз.* Важно, что карта рестрикции может быть получена для любой последовательности ДНК, *вне зависимости от того,* знаем мы что-нибудь про функции данной последовательности или нет.

3.4 Организация прерывистых генов консервативна

Ключевые понятия:

- С помощью рестрикционных карт или электронной микроскопии можно провести сравнение генов и их РНК-продуктов: отсутствующие в зрелых мРНК регионы генов и будут являться интронами.
- Сравнение гомологичных генов из разных организмов позволяет сделать вывод о консервативности расположения интронов, однако длины аналогичных (по положению в гене) интронов могут значительно варьировать.
- Обычно интроны не кодируют белки.

Рестрикционная карта безынтронного гена в точности соответствует карте его мРНК.

Если ген содержит интроны, карта концов этого гена соответствует карте концов мРНК, а в пределах гена карты могут различаться, поскольку дополнительные регионы, содержащиеся в гене, не представлены в мРНК. Каждый такой регион соответствует интрону. На РИС. 3.5 показано сравнение рестрикционных карт гена β-глобина и его мРНК. В приведенном примере ген содержит два интрона, каждый из которых имеет несколько сайтов рестрикции, отсутствующих в так называемой комплементарной ДНК (кДНК). Набор сайтов рестрикции в экзонах совпадает и в кДНК, и в гене.

Таким образом, сравнение нуклеотидной последовательности геномной ДНК и мРНК позволяет точно определить расположение интронов. Как видно из РИС. 3.6, в интронах обычно нет открытых рамок считывания. За

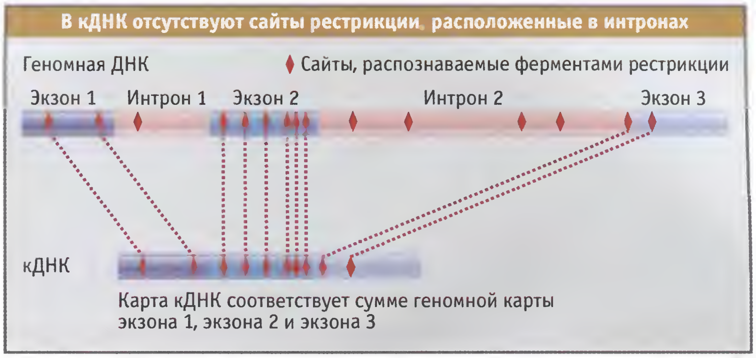


РИС. 3.5 **Сравнение карт рестрикции кДНК и геномной ДНК гена β глобина мыши позволяет увидеть, что ген содержит два интрона, которых нет в кДНК. Экзоны кДНК и геномной ДНК полностью соответствуют и по длине и по порядку расположения**

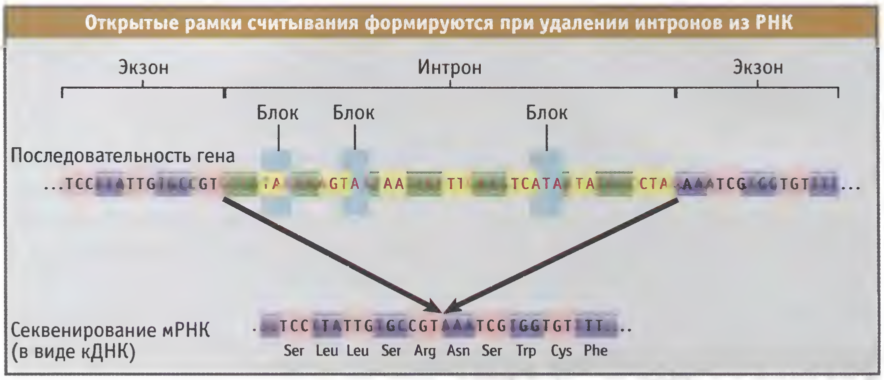


РИС. 3.6 **Интрон — это последовательность, присутствующая в гене и отсутствующая в мРНК. Рамка считывания выделена светлыми и затененными блоками; обратите внимание, что все три рамки считывания в интроне блокированы стоп-кодонами**

счет удаления интронов в мРНК формируется единая непрерывная рамка считывания.

Структура эукариотических генов демонстрирует значительную изменчивость. Некоторые гены являются безынтронными, и их геномная последовательность колинеарна последовательности мРНК. Большая же часть генов высших эукариот имеет прерывистую структуру, а число и размеры интронов в них сильно варьируют.

Все классы генов могут быть прерывистыми: кодирующие белки, кодирующие рРНК и кодирующие тРНК. Интроны найдены также в митохондриальных генах низших эукариот и в генах хлоропластов. Прерывистые гены не являются исключительной чертой эукариот, они также найдены у бактерий и бактериофагов. Однако у прокариот они встречаются крайне редко.

Хорошо изученным примером прерывистых генов являются гены глобинов (см. разд, 3.10 «Гены, относящиеся к одному семейству, имеют схожую организацию»). Два основных типа глобиновых генов, α и β, имеют схожую структуру. О постоянстве организации глобиновых генов у млекопитающих свидетельствует структура «общего» глобинового гена, приведенного на РИС. 3.7.

Промежутки в кодирующей белок последовательности всех известных глобиновых генов млекопитающих, птиц и лягушек расположены в одних и тех же местах. Первый интрон, как правило, короче второго (хотя их величина может варьировать). Наибольшее разнообразие в длину глобиновых генов вносит вариабельность второго интрона. У мышей второй интрон α-глобинового гена имеет длину всего в 150 пн, при этом длина всего гена достигает 850 пн, тогда как в β-глобиновом гене увеличение длины интрона до 585 пн приводит к увеличению общей длины до 1382 пн. Вариация длин генов значительно выше, чем вариация длин мРНК (длина мРНК α-глобина составляет 585 пн, а β-глобина — 620 пн).

На РИС. 3.8 приведен пример со сравнительно длинным геном дигидрофолатредуктазы (DHFR). У млекопитающих ген DHFR содержит шесть экзонов, соответствующих 2000 оснований мРНК. Но за счет протяженных интронов длина его ДНК значительно больше. У трех млекопитающих экзоны гена DHFR в высшей степени сходны, и относительное расположение интронов также демонстрирует сходство. При этом длина отдельных интронов сильно варьирует, что приводит к изменению длины гена от 25 до 31 тпн.

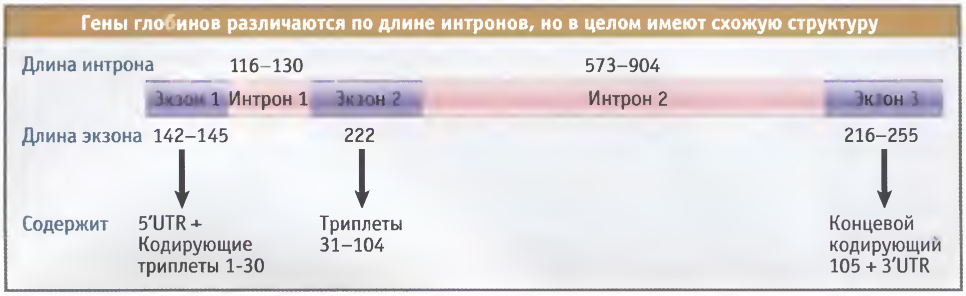


РИС. 3.7 Все функциональные гены глобинов имеют прерванную структуру, характеризующуюся наличием трех экзонов. Длины, приведенные на рисунке, соответствуют генам β-глобина млекопитающих (UTR-нетранслируемый регион)

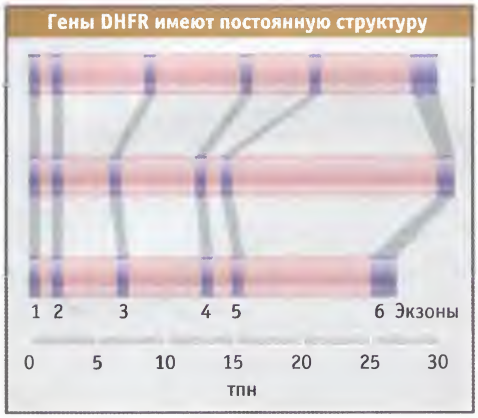


РИС. 3.8 Гены DHFR млекопитающих имеют схожую структуру, характеризующуюся короткими экзонами и очень длинными интронами, однако по длине интронов они сильно отличаются

Примеры с генами глобинов и DHFR иллюстрируют следующий феномен: *эволюционно родственные гены имеют схожую организацию, демонстрирующую постоянство расположения (по крайней мере, некоторых) интронов. Различия же в длинах генов в основном определяются длиной интронов.*

3.5 Последовательности экзонов консервативны, а интронов — вариабельны

Ключевые понятия:

- При межвидовом сравнении родственных генов видно, что последовательности экзонов консервативны, тогда как интроны не похожи.
- Интроны эволюционируют быстрее, чем экзоны. Отбор воздействует на интроны в значительно меньшей степени, поскольку большая часть их последовательности не влияет на функциональность белка.

Уникальна ли структура гена для данного генома? На этот вопрос можно ответить двояко. С одной стороны, размер конкретного гена уникален, но, с другой стороны, его экзоны могут быть дублированы экзонами других генов. Как правило, если два гена гомологичны, то сходство последовательностей их экзонов выше, чем их интронов. Существуют примеры, когда экзоны двух генов кодируют один и тот же белок, в то время как их интроны различаются. Скорее всего, два таких гена возникли в результате дупликации общего предкового гена, а различия накопились позже, однако мутации не затронули экзоны, поскольку на них лежит ответственность за синтез функционального белка.

Позже мы увидим, что при рассмотрении эволюции гена экзоны можно считать элементарными кирпичиками, собранными в различных комбинациях. Часть экзонов гена может совпадать с экзонами другого гена, при этом остальные их экзоны будут различны. Как правило, для таких случаев интроны оказываются вовсе несопоставимы по последовательности. Такие гены могут возникать посредством дупликации или транслокации отдельных экзонов.

Сходство последовательностей двух генов можно представить в виде точечной диаграммы. Подобное сравнение приведено на РИС. 3.9. Точки обозначают нуклеотидные позиции, по которым совпадают анализируемые последовательности. Если последовательности идентичны, точки формируют линию, проходящую под углом 45°. В то же время линия прерывается на несовпадающих участках или сдвигается в вертикальном или горизонтальном направлениях (в местах делеций и инсерций), по которым различаются анализируемые последовательности.

При сравнении генов β-глобинов мыши такая линия проходит через три экзона и один короткий интрон. Линия прерывается во фланкирующих регионах и в районе большого интрона. Это достаточно типичная картина — кодирующие последовательности и прилегающие к ним участки демонстрируют значительное сходство, тогда как в местах расположения длинных интронов и на концевых участках гена последовательности отличаются.

Степень различий между двумя экзонами связана с различиями между белками, кодируемыми этими экзонами. Чаще всего различия гомологичных эк-

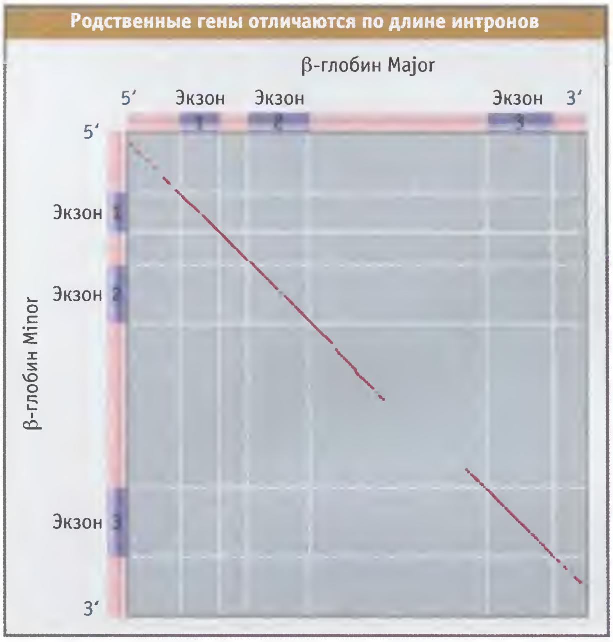


РИС. 3.9 Последовательности генов β^{maj}- и β^{min}-глобина мышей схожи структурой кодирующих участков, однако отличаются фланкирующими участками и последовательностью большого интрона. Данные предоставлены Филипом Лидером из Гарвардской медицинской школы

зонов представлены однонуклеотидными заменами. Поскольку экзоны, представляющие собой транслируемые регионы гена, кодируют белок, потенциал к изменению их нуклеотидной поледовательности значительно меньше, чем для нетранслируемых регионов гена. Некоторые нуклеотидные замены не влияют на значение кодона, поскольку измененный кодон может соответствовать той же аминокислоте. Гораздо чаще изменения появляются в нетранслируемых регионах (соответствующих 5′- и 3′-концевым участкам мРНК).

Соответствующие интроны могут отличаться по двум параметрам: длине (из-за инсерций и делеций) и последовательности нуклеотидов (из-за замен оснований). Эволюция влияет на интроны значительно сильнее, чем на кодирующие белок регионы гена. При сравнении генов у представителей различных видов часто видно, что экзоны одного и того же гена гомологичны, а интроны различаются настолько, что сходство между соответствующими последовательностями практически отсутствует.

И в экзонах, и в интронах частота возникновения мутаций одинакова, но из-за действия отрицательного отбора изменения экзонов более эффективно удаляются из популяции. Кроме того, в результате отсутствия ограничений, вызванных необходимостью кодировать белки, интроны довольно легко накапливают однонуклеотидные замены и другие изменения. Это доказывает, что последовательность интронов не является столь уж важной для наработки кодируемого белка. Более того, неясно, влияет ли их присутствие каким-либо образом на функцию гена.

3.6 Длина генов сильно варьирует

Ключевые понятия:

- У дрожжей многие гены непрерывны (не содержат интронов), тогда как у высших эукариот большинство генов имеют прерывистое строение.
- Обычно экзоны короче интронов и кодируют не более 100 аминокислот.
- У низших эукариот интроны сравнительно короткие, тогда как у высших эукариот их длина может достигать десятков тысяч пар нуклеотидов.
- Длина эукариотического гена в основном определяется длиной его интронов.

На РИС. 3.10 показана общая структура организации генов у дрожжей, насекомых и млекопитающих. У *Saccharomyces cerevisiae* подавляющее число генов (>96%) не содержат интронов, а интронсодержащие гены довольно компактны. У *S. cerevisiae* практически нет генов, содержащих более четырех экзонов.

У насекомых и млекопитающих обратная ситуация. У них лишь немногие гены имеют непрерывную структуру (у млекопитающих — около 6%). В генах насекомых, как правило, мало экзонов, редко более 10. Гены млекопитающих разбиты на большее число фрагментов (некоторые гены содержат десятки экзонов). Приблизительно половина генов млекопитающих несет более 10 интронов.

Оценивая последствия подобной организации для размера генов, мы увидим, что различия между дрожжами и высшими эукариотами весьма значительны (РИС. 3.11). Средняя длина гена у дрожжей 1,4 тпн, и только немногие из них длиннее 5 тпн. Превалирование прерывистых генов у эукариот означает, что их гены намного длиннее кодирующих белок частей. Только некоторые гены мух и млекопитающих короче 2 тпн. Длина многих из них варьирует в диапазоне от 5 до 100 тпн. В среднем длина человеческого гена составляет 27 тпн (см. рис. 5.11).

Переключение с преимущественно безынтронных генов на преимущественно прерывистые происходит на уровне низших эукариот. У грибов (за исключением дрожжей) большинство генов уже имеет прерывистую структуру, однако содержит сравнительно малое число экзонов (<6) и имеет небольшую длину (<5 тпн.) Переход к длинным генам происходит у высших эукариот. Их размер становится заметно больше уже у насекомых. С увеличением длины гена корреляция между размером генома и сложностью организации организма исчезает (см. рис. 4.5).

С увеличением размера генома интроны демонстрируют тенденцию к увеличению, при том что экзоны остаются довольно короткими.

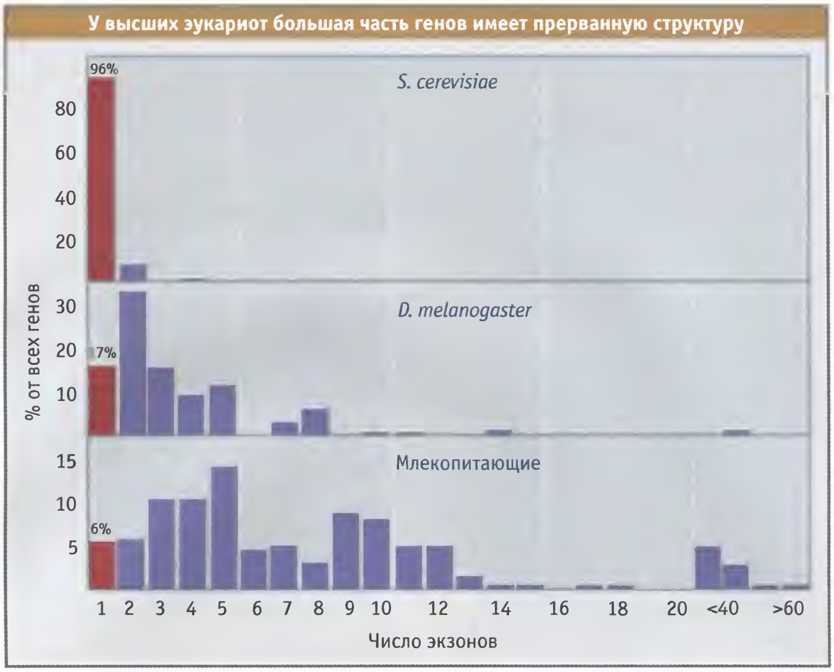


РИС. 3.10 У дрожжей большая часть генов имеет непрерывную структуру, в то время как у мух и млекопитающих большинство генов прерывистые. (Непрерывные гены содержат только один экзон и собраны в колонке, расположенной на графике слева.)

На **РИС. 3.12** показано, что экзоны в ряду рассматриваемых организмов становятся короче. У высших эукариот экзон среднего размера кодирует около 50 аминокислот, и общее распределение хорошо отражает идею о том, что гены эволюционировали путем добавления участков, кодирующих небольшие отдельные домены белков (см. разд. 3.8 «Как эволюционировали прерывистые гены?»). У разных групп высших эукариот размеры экзонов остаются практически неизменными. Тем не менее у позвоночных они более компактны (лишь немногие экзоны длиннее 200 пн), а кодирующие последовательности у дрожжей представляют собой один длинный непрерывающийся экзон. Для экзонов, кодирующих нетранслируемые 5'- и 3'-регионы, отмечена тенденция увеличения размеров относительно экзонов, кодирующих собственно аминокислотную последовательность.

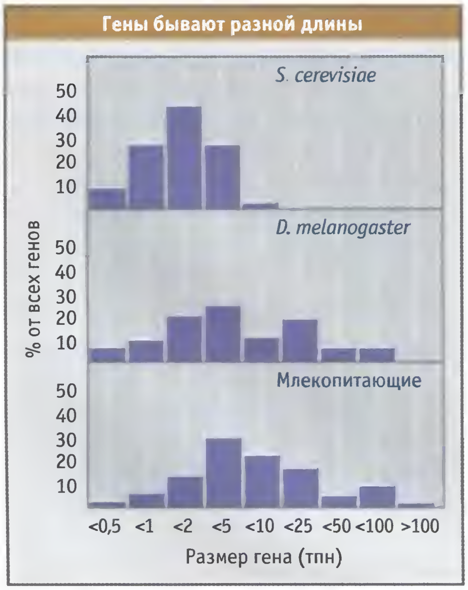


РИС. 3.11 У дрожжей гены преимущественно короткие, а у мух и млекопитающих они варьируют в широком диапазоне — от коротких до очень длинных

На **РИС. 3.13** показано, что интроны могут сильно отличаться по размеру. У червей и мух средний интрон практически равен экзону. Однако у червей почти нет длинных интронов, а у мух они уже составляют заметную часть. У позвоночных размеры интронов представлены в более широком диапазоне: от сравнимых по размеру с экзонами (<200 пн) до длин, измеряемых десятками тпн и достигающих в отдельных случаях 50–60 тпн.

Как правило, очень длинные гены своим размером обязаны именно интронам, и их длина никак не связана с величиной кодируемого продукта. Также у высших эукариот отсутствует корреляция между размером гена, длиной мРНК и количеством экзонов. Размер гена, таким образом, в первую очередь зависит от длины его интронов. У млекопитающих, насекомых и птиц ген среднего размера примерно в 5 раз длиннее своей мРНК.

3.7 Некоторые последовательности ДНК кодируют более одного белка

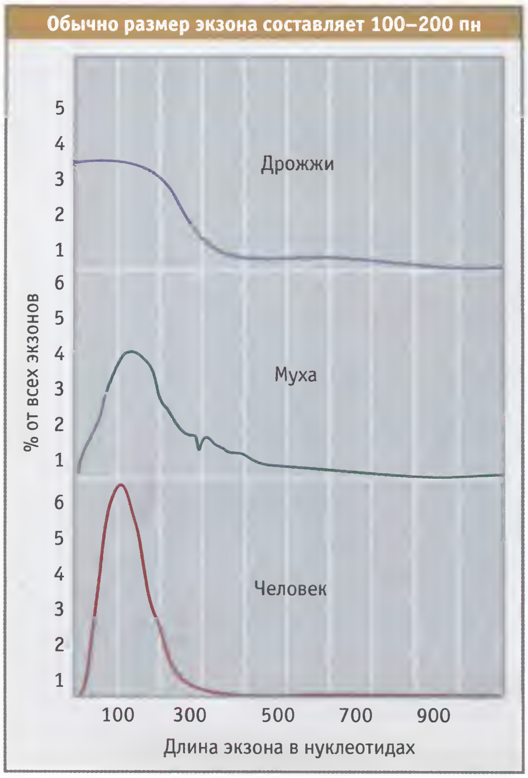


РИС. 3.12 Экзоны, кодирующие белки, обычно короткие

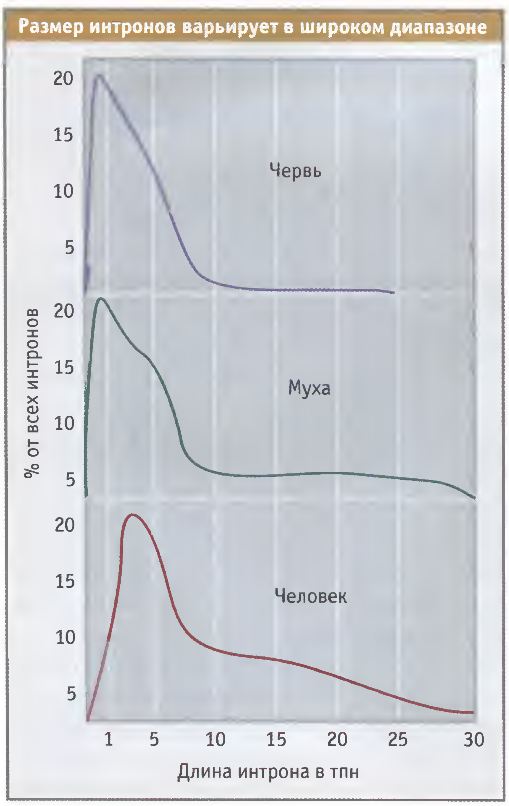


РИС. 3.13 Интроны могут быть как короткими, так и очень длинными

Ключевые понятия:

- Возможность использования дополнительных кодонов, инициирующих и терминирующих транскрипцию, позволяет получить с одного гена два белка, один из которых соответствует части другого.
- Одна и та же последовательность ДНК может отвечать за синтез негомологичных белков при условии, что информация считывается в разных рамках для двух перекрывающихся генов.
- Гомологичные белки, различающиеся наличием или отсутствием определенных регионов, могут возникать в результате альтернативного сплайсинга. Это может быть включение или исключение из зрелой РНК отдельных экзонов (или выбора между двумя альтернативными экзонами).

Большинство генов состоит из последовательности ДНК, кодирующей единственный белок (несмотря на то, что ген может содержать интроны и некодирующие концевые участки). Тем не менее в некоторых случаях один и тот же участок ДНК может кодировать более одного белка.

Перекрывающиеся гены возникают, когда некий ген одновременно является частью другого, меньшего гена. Начало или конец такого гена кроме кодирования полноразмерного белка может одновременно кодировать укороченную форму белка (**РИС. 3.14**). Результат такого перекрывания кодирующих регионов обычно выглядит так же, как если бы произошло расщепление части молекул полноразмерного белка с образованием двух форм — полноразмерной и неполной.

Более сложный случай перекрывания — когда одна и та же последовательность ДНК принимает участие в синтезе двух *негомологичных* белков. В этой ситуации одна последовательность ДНК считывается более чем в одной рамке. В генах клеточных организмов ДНК обычно считывается лишь в одной из трех возможных рамок считывания. Между тем, в некоторых вирусных и митохондриальных генах два прилегающих друг к другу гена могут перекрываться, транслируясь в различных рамках считывания. Пример такого перекрывания приведен на **РИС. 3.15**. Обычно размер перекрывания сравнительно невелик, так что большая часть последовательности, кодирующей каждый из белков, представлена отдельным регионом ДНК.

В некоторых генах система *альтернативного сплайсинга* переключает процесс соединения экзонов. Таким образом, один ген может производить разные молекулы мРНК, отличающиеся набором экзонов. Присутствие некоторых экзонов в таких мРНК может быть необязательным (то есть они могут как присутствовать, так и отсутствовать в зрелой мРНК). Кроме того, некоторые экзоны могут быть взаимоисключающими (в зрелой мРНК может присутствовать лишь один из двух экзонов). Альтернативные формы мРНК производят белки, в которых одна часть общая, а другая вариабельная.

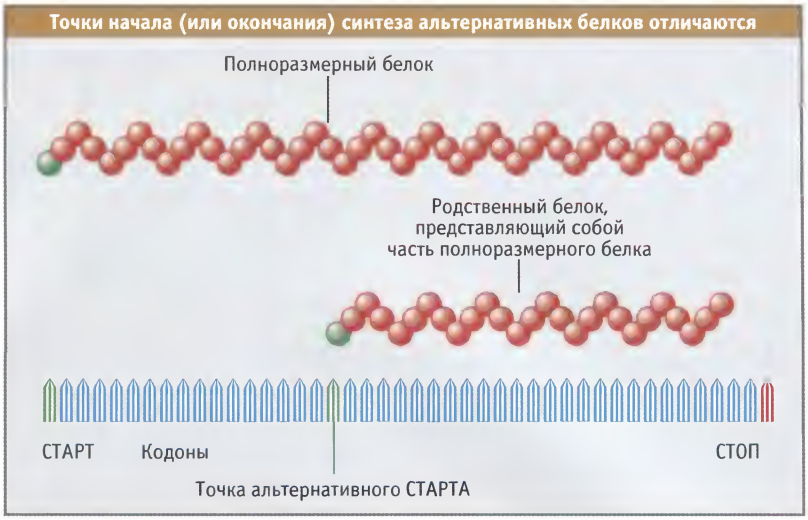


РИС. 3.14 Два белка могут быть синтезированы с одного гена путем инициации (или терминации) экспрессии в разных точках



РИС. 3.15 Одна и та же последовательность может использоваться двумя генами, но считывание такой последовательности будет идти в различных рамках

В некоторых случаях альтернативный сплайсинг не влияет на последовательность белка. Например, изменения, затрагивающие 5′- или 3′-нетранслируемые регионы, могут иметь регуляторные последствия, но белок будет производиться один. Также существуют случаи, когда один экзон заменяется другим (РИС. 3.16).

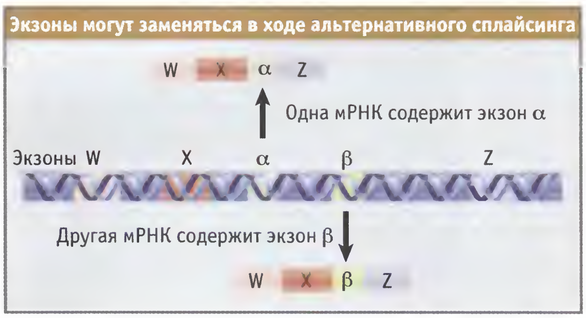


РИС 3.16 Альтернативный сплайсинг приводит к формированию α- и β-вариантов тропонина T

В приведенном примере белки, кодируемые двумя альтернативными мРНК, содержат одинаковый протяженный регион, но различаются регионом альтернативного сплайсинга. 3′-часть гена тропонина T, кодирующего мышечный белок крыс, содержит пять экзонов, но для построения мРНК используются только четыре из них. Три экзона (W,X и Z) присутствуют в обоих вариантах экспрессии гена. Однако в одном из вариантов между X и Z присутствует экзон α, а в другом — β. Таким образом, α- и β-формы тропонина T отличаются последовательностью аминокислот на участке между экзонами X и Z и синтезируются в зависимости от того, какой из альтернативных экзонов (α или β) был использован. Любой из двух альтернативных экзонов может быть использован для синтеза конкретной мРНК, но они оба в одной и той же зрелой мРНК не встречаются.

На РИС. 3.17 приведен пример, в котором в результате альтернативного сплайсинга один из экзонов включается лишь в некоторые зрелые мРНК. Ген продуцирует единственный тип транскриптов, однако их сплайсинг

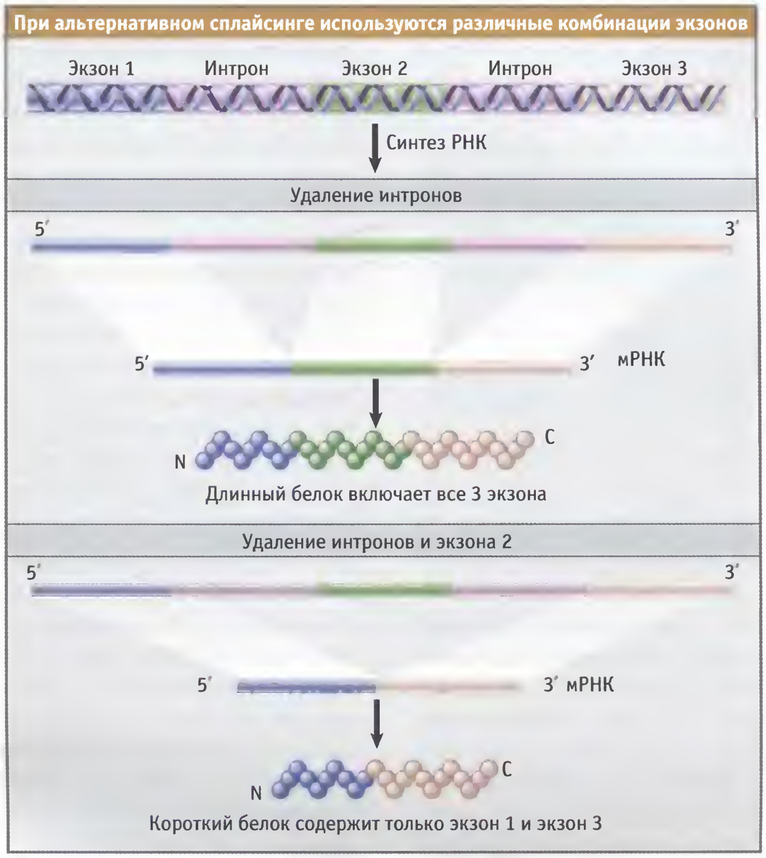


РИС. 3.17 Для формирования мРНК с различными комбинациями экзонов при альтернативном сплайсинге используется одна и та же пре-мРНК

может осуществляться двумя путями. В первом случае два интрона сплайсируются, и три экзона соединяются вместе. В другом случае второй экзон не распознается. В результате сплайсируется как бы один большой интрон. Этот интрон состоит из интрона 1, экзона 2, и интрона 2. Второй экзон в данном случае не был воспринят системой сплайсинга как самостоятельный элемент гена. Два пути сплайсинга дают два разных белка. Они одинаковы по концам, но один из белков имеет дополнительную последовательность посередине. Таким образом, данный регион ДНК кодирует более чем один белок (другие типы альтернативного сплайсинга обсуждаются в разделе 26.12 «Альтернативный сплайсинг — это реализация различных вариантов сшивания экзонов между собой»).

Иногда могут быть задействованы одновременно оба описанных варианта альтернативного сплайсинга. В этом случае по каждому из путей синтезируются определенные пропорции мРНК. Иногда для разных типов клеток или при разных условиях можно наблюдать различающиеся пути сплайсинга.

Таким образом, альтернативный (или дифференциальный) сплайсинг ведет к синтезу с одного и того же гена белков с частично разными последовательностями аминокислот. Интересно, что большие геномы эукариот с протяженными (рассредоточенными) генами вдобавок обладают способностью к синтезу различных продуктов с одного локуса. У мух и червей альтернативный сплайсинг увеличивает количество белков относительно числа генов примерно на 15%, в то время как у человека альтернативно экспрессируется около 60% генов (см. разд. 5.5 «В геноме меньше генов, чем кажется»). Около 80% случаев альтернативного сплайсинга приводит к изменению белковой последовательности.

3.8 Как эволюционируют прерывистые гены?

Ключевые понятия:

- Основной вопрос эволюции: «Как возникли гены?» Имели они прерывистую структуру с самого начала или возникли как непрерывные последовательности и затем были «обогащены» интронами?
- Вероятно, подавляющая часть генов, кодирующих белки, возникла в виде прерывистой формы, однако гены, кодирующие рибосомальные и транспортные РНК, изначально могли быть непрерывны.
- Существует особый класс мобильных интронов, способных самостоятельно встраиваться в гены.

Прерывистая структура эукариотических генов позволяет представить эукариотический геном в виде «моря

интронов» (часто неуникальных по своей последовательности), в котором встречаются островки экзонов (иногда очень коротких). Эти островки собраны в отдельные архипелаги, представляющие собой гены.

Как же изначально выглядели прерывистые гены?

- В соответствии с моделью «ранних интронов» интроны всегда были неотъемлемой частью генов. Гены возникли как прерывистые структуры, а существующие теперь безынтронные гены утеряли интроны в процессе эволюции.
- Модель «поздних интронов», напротив, предполагает, что предковые кодирующие единицы не содержали интронов, а интроны были «вставлены» в них позже.

Чтобы выбрать более верную модель, необходимо выяснить, чем объясняются различия между генами эукариот и прокариот. Вызваны они приобретением интронов эукариотами или же их потерей прокариотами?

Модель «ранних интронов» предполагает, что мозаичная структура гена является древним подходом к изменению структуры гена с целью эффективного создания новых белков. Можно предположить, что древняя клетка содержала набор генов с экзон-интронной структурой, а основным аспектом их эволюции была перетасовка экзонов с целью построения новых белков.

Если представить себе перетасовку регионов безинтронных генов, очевидно, что каждая такая операция требует исключительно точной рекомбинации молекул ДНК с сохранением рамки считывания и в той же ориентации. Кроме того, если рекомбинация оказалась неудачной, клетка теряет исходные кодирующие единицы и должна погибнуть. Если же предположить, что в результате рекомбинации две «структурных единицы» белка (например, два домена, каждый из которых закодирован своим экзоном) оказались в пределах одной единицы транскрипции, то для создания нового эволюционно выгодного варианта белка достаточно провести «проверку» множества вариантов альтернативного сплайсинга. Если получающиеся варианты «эволюционно неинтересны», то исходные единицы (экзоны) остаются доступными для последующих рекомбинаций. Подход с альтернативным сплайсингом позволяет клетке как бы тестировать белок на устойчивость к удалению или добавлению той или иной его части, не причиняя при этом вреда ДНК. Описанный механизм подтверждается тем фактом, что схожие экзоны могут располагаться в различных генах так, как будто бы гены были собраны путем перетасовки экзонов (см. разд. 3.9 «Некоторые экзоны напоминают последовательности, кодирующие целые белки»).

На РИС. 3.18 приведен пример переноса последовательности, включающей экзон, на новое место в геноме. Поскольку экзоны, как правило, значительно короче интронов, выглядит логичным, что при рекомбинации добавляемый регион встроится в интрон, а не в экзон. Для сплайсинга необходимы только последовательности, расположенные в местах стыка экзонов и интронов (экзон фланкирован функциональными участками). При сплайсинге распознаются пары близкорасположенных сайтов сплайсинга: 5′-соединение интрона с экзоном взаимодействует с другим (3′) концевым стыком этого интрона и последующего экзона. В случае добавления в структуру гена нового экзона (рис. 3.18) его фланги будут работать

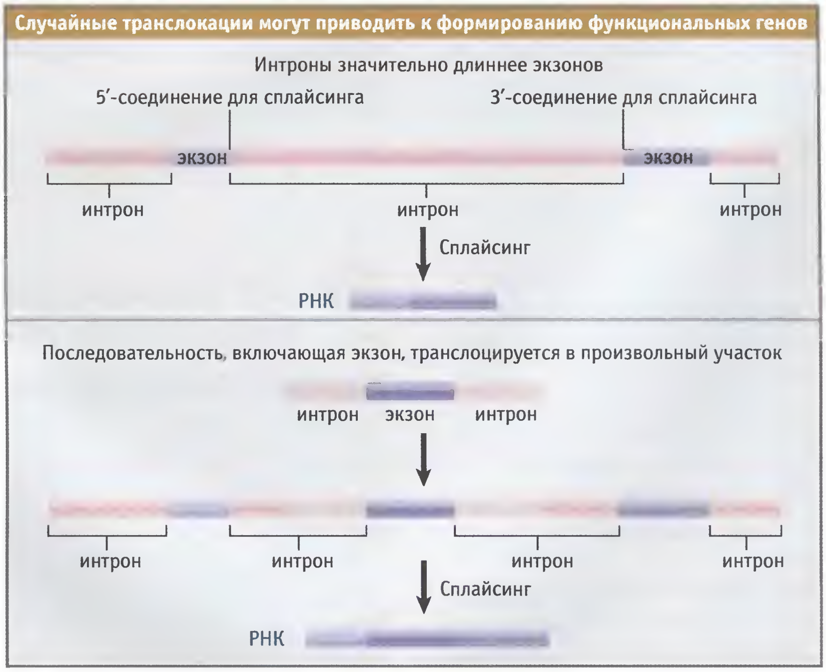


РИС. 3.18 При транслокации в интрон экзона, маркированного фланкирующими последовательностями, он может быть включен в РНК

при сплайсинге так же, как и фланги уже существовавших экзонов. В результате, вновь добавленный экзон встроится в мРНК между двумя исходными экзонами. При этом новый экзон автоматически окажется в той же рамке считывания, что и старые, и с него будет синтезирован новый регион белка. Описанный механизм мог быть ответствен за создание новых комбинаций экзонов в процессе эволюции. Обратите внимание, что приведенный механизм напоминает метод улавливания экзонов, используемый для поиска функциональных экзонов (рис. 4.11).

Иногда можно обнаружить альтернативные формы генов рРНК и тРНК (с интронами и без). В случае с тРНК, для которых все молекулы обладают единой пространственной структурой, маловероятно, что в эволюции они независимо формировались за счет соединения нескольких частей генов. Кроме того, разные регионы тРНК должны быть комплементарны (и формировать двухцепочечные структуры), что крайне важно для поддержания пространственной структуры молекулы. То есть в данном случае интроны должны были вставляться в уже существующие безинтронные гены.

Геномы органелл являются как бы связующим звеном между прокариотами и эукариотами. Существует множество сходных черт между митохондриями и хлоропластами с одной стороны и бактериями с другой, и это позволяет предположить, что органеллы произошли как следствие *эндосимбиоза*, в ходе которого древняя бактерия попала в «эукариотическую» цитоплазму. Однако сходству органелл с бактериями в процессах синтеза РНК и белков можно противопоставить наличие в генах органелл интронов. Этим они напоминают скорее эукариотические ядерные гены.

Интроны найдены в нескольких генах хлоропластов, включая и те из них, которые гомологичны генам *E. coli*. То есть можно предположить, что эндосимбиоз произошел до того, как в прокариотической линии исчезли интроны. Для некоторых генов путем сравнения гомологов можно проследить эволюцию до момента, когда произошел эндосимбиоз.

Митохондрии демонстрируют еще более поразительный пример. Митохондриальные геномы дрожжей и млекопитающих, хотя и существенно отличаются по своей организации, кодируют практически идентичные белки. Митохондриальные геномы позвоночных очень малы, и их непрерывные гены организованы очень компактно, в то время как митохондриальные геномы дрожжей значительно больше по размеру и демонстрируют наличие сложных прерывающихся генов. Какой же из этих геномов является предковым? Интроны митохондриальных (и некоторых других) генов дрожжей обладают способностью к перемещению — они представляют собой последовательности, способные вырезаться из РНК и встраиваться в виде ДНК-копий в новые места генома, что позволяет предположить их «вторичное» происхождение (см. разд. 27.5 «Некоторые интроны группы I кодируют эндонуклеазы, поддерживающие их мобильность» и разд. 27.6 «Интроны группы II могут кодировать полифункциональные белки»).

3.9 Некоторые экзоны напоминают последовательности, кодирующие целые белки

Ключевые понятия:

- В пользу того, что экзоны были строительным материалом эволюции и что первые гены имели прерывистую структуру говорят следующие факты:
  - у довольно сильно отстоящих друг от друга видов гены имеют схожую экзон-интронную структуру;
  - существует большое число экзонов, последовательности которых могут быть сопоставлены с существующими последовательностями, кодирующими целые белки;
  - существует гомология между экзонами, находящимися в разных генах.

Если современные белки эволюционировали путем комбинирования исходно независимых коротких предковых белков, то прибавление единиц, скорее всего, происходило последовательно на протяжении некоторого времени. Можно ли в современных белках обнаружить деление составляющих их единиц по функциям? Иными словами, можем ли мы сопоставить отдельные функции нынешних белков с отдельными экзонами?

Действительно, в некоторых случаях есть четкая взаимосвязь между структурами гена и белка. Хорошим примером являются иммуноглобулины, кодируемые генами, у которых каждый экзон в точности соответствует определенному функциональному домену белка. На РИС. 3.19 дано сравнение структуры иммуноглобулина и его гена.

Иммуноглобулин — это тетрамер, состоящий из двух легких и двух тяжелых цепей, которые, объединяясь, формируют белок с несколькими различными доменами. Легкие и тяжелые цепи отличаются по структуре, при этом существует несколько вариантов тяжелых цепей. Каждая из цепей синтезируется с гена, содержащего серию экзонов, соответствующих структурным доменам белка.

В ряде других случаев отдельным экзонам гена могут быть приписаны отдельные функции. У секреторных белков первый экзон, кодирующий N-концевой участок полипептида, часто определяет сигнальную последовательность, вовлеченную в процесс секреции. В качестве примера можно привести инсулин.

Представление об экзонах как о функциональных единицах генов подтверждается примерами, в которых два гена содержат гомологичный экзон (экзоны), в то время как остальные их экзоны непохожи. На РИС. 3.20 проиллюстрировано сходство между рецептором липопротеинов низкой плотности плазмы (ЛПНП) и другими белками человека. В центре гена рецептора липопротеинов низкой плотности расположена серия экзонов, гомологичных гену предшественника эпидермального фактора роста (EGF). В N-концевой части белка группа экзонов кодирует последовательность, гомологичную белку комплемента C9. Таким образом, можно предположить, что

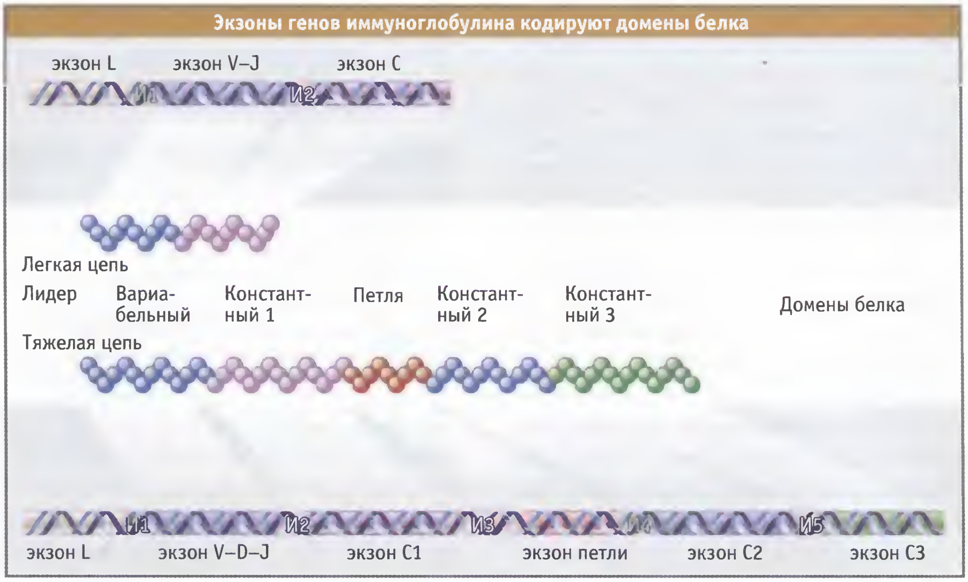


РИС. 3.19 **Легкие и тяжелые цепи иммуноглобулина закодированы генами, которые по экзон-интронной структуре соответствуют отдельным доменам белка. Каждый домен белка соответствует экзону; интроны пронумерованы цифрами от 1 до 5**

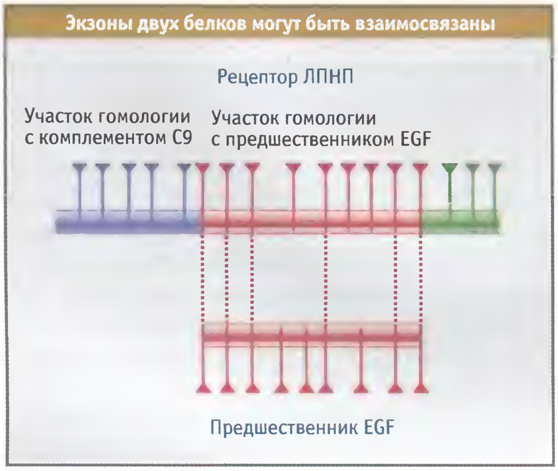


РИС. 3.20 **Ген рецептора ЛПНП включает 18 экзонов, некоторые из которых гомологичны экзонам предшественника EGF, а некоторые — генам комплемента крови C9. Треугольниками отмечены позиции интронов**

ген рецептора липопротеинов низкой плотности возник в результате сборки *модулей*, ответственных за различные функции. Эти модули в разных комбинациях используются и в других белках.

Экзоны, в большинстве своем, довольно малы (см. рис. 4.11) и по размеру примерно соответствуют самому маленькому полипептиду, способному сохранять стабильную конформацию (около 20–40 аминокислот). Возможно, изначально белки собирались из еще меньших модулей. Модуль не обязательно соответствует отдельной функции современного белка; несколько модулей могли объединяться и формировать новую функцию. С увеличением длины белка число кодирующих его экзонов, как правило, увеличивается. Этот факт соотносится с представлением о том, что белки становятся полифункциональными путем успешного добавления подходящего модуля.

Высказанные предположения могут объяснить и другую общую черту современных белков: участки, соответствующие границам экзонов и интронов, часто расположены на поверхности белка. Поскольку в свете высказанной гипотезы белки собирались модульно, связи (по крайней мере, последних добавленных модулей) демонстрируют тенденцию к расположению на поверхности.

3.10 Гены, относящиеся к одному семейству, имеют схожую организацию

Ключевые понятия:

- Схожесть черт у генов, относящихся к одной группе, позволяет сделать предположение о функции, выполняемой ими до разделения в эволюции.
- Все гены глобинов схожи по своей структуре. Они содержат три экзона и два интрона, что дает основания предполагать, что они произошли от одного предкового гена.

Внутри эукариотического генома множество генов гомологичны друг другу (формируя семейство генов).

Семейство генов — это группа генов, кодирующих родственные или идентичные белки. Семейство возникает путем дупликации гена. Поначалу две копии гена идентичны, однако позднее в результате мутаций они накапливают различия. Последующие дупликации и отклонения от исходной структуры приводят к дальнейшему увеличению семейства. Гены глобинов представляют собой пример семейства, которое может быть разделено на два подсемейства (α-глобинов и β-глобинов). Члены подсемейств имеют схожее строение и выполняют одинаковые функции. Предложенную классификацию можно распространить и на случаи с генами, обладающими значительно меньшим сходством, но, тем не менее, имеющим черты, позволяющие судить об их общем происхождении. В этих случаях группы семейств генов могут формировать **суперсемейства**.

Прекрасным примером эволюционной стабильности являются α- и β-глобины и гомологичные им белки. Аминокислотная последовательность миоглобина — мономерного кислород-связывающего белка животных — позволяет предположить его происхождение из глобиновой субъединицы. Леггемоглобин, так же как и миоглобин, представляет собой мономерный кислородсвязывающий белок, но не животных, а бобовых растений, и тоже делит общее происхождение с другими гем-связывающими белками. Вместе α- и β-глобины, миоглобин и леггемоглобин составляют суперсемейство глобинов — группу из нескольких семейств генов, объединенных происхождением от общего предшественника.

И α-, и β-глобиновые гены содержат три экзона (рис. 3.7). Два интрона имеют постоянное расположение относительно кодирующей последовательности, а центральный экзон кодирует гем-связывающий домен глобиновой цепи.

Миоглобин представлен единственным геном в геноме человека, и его структура черезвычайно похожа на структуру глобиновых генов. Таким образом, трехэкзонная структура возникла в эволюциии этих генов до того, как произошло разделение их функций.

Гены леггемоглобинов содержат три интрона, первый и последний из которых расположены в местах кодирующей последовательности, голомологичных местам расположения двух интронов глобиновых генов. Это поразительное сходство указывает на очень древнее происхождение гем-связывающих белков от общего предкового гена (РИС. 3.21).

Центральный интрон леггемоглобина разделяет два экзона, вместе кодирующих ту же последовательность, что и один центральный экзон глобинов. Мог ли центральный экзон глобинового гена получиться в результате соединения двух центральных экзонов предкового гена? Или же предковой формой является структура с одним центральным экзоном, и в этом случае интрон был вставлен в последовательность в начале эволюции растений?

Случаи, в которых гомологичные гены отличаются по структуре, могут нести информацию об их эволюции. В качестве еще одного примера можно привести

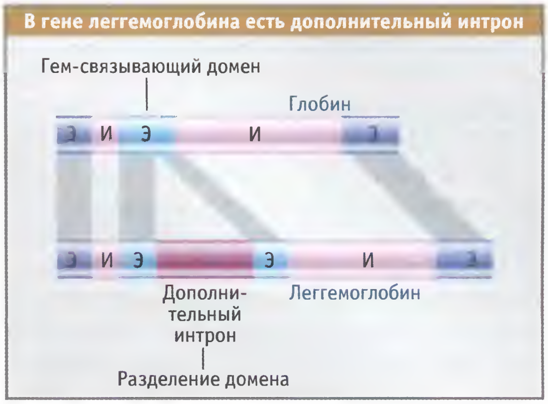


РИС. 3.21 Структура гем-связывающего экзона генов глобина ответственна за функцию белка, однако ген леггемоглобина несет в центральной части этого домена дополнительный интрон

инсулин. У млекопитающих и птиц инсулин кодирует только один ген (за исключением грызунов, у которых два таких гена). На РИС. 3.22 изображена структура этих генов.

Основной принцип, который мы используем при сравнении организации родственных генов у различных видов, состоит в том, что *общие черты идентифицируют структуру, предшествовавшую разделению двух вариантов в эволюции*. У кур единственный инсулиновый ген содержит два интрона; такую же структуру имеет один из двух родственных генов крыс. Сходство структур указывает на то, что предковый ген содержал два интрона. Однако во втором гене у крыс только один интрон. Вероятнее всего, этот вариант гена произошел у грызунов в результате дупликации и последующего удаления одного из интронов.

Организация некоторых генов демонстрирует значительные отличия у представителей разных видов. Эти



РИС. 3.22 Эволюция гена инсулина крыс шла по пути удаления интрона из гена-предшественника, несущего два интрона

случаи должны объясняться многократными вставками и удалениями интронов в процессе эволюции.

Хорошо охарактеризованным примером, иллюстрирующим высказанное предположение, являются гены актинов. Типичный актиновый ген состоит из 5′-нетранслируемого региона (протяженностью менее 100 нуклеотидов), кодирующего региона (длиной около 1200 нуклеотидов) и 3′-нетранслируемого региона (около 200 нуклеотидов). Большинство актиновых генов имеет прерывистую структуру, при этом большинство интронов расположены в пределах кодирующей последовательности (за исключением одного интрона, иногда расположенного в 5′-нетранслируемом регионе).

На **РИС. 3.23** показано, что почти все гены актинов отличаются по локализации интронов. Если их сравнить, станет заметно, что интроны могут располагаться в 19 различных местах. Однако гены, содержащие более шести интронов, отсутствуют; некоторые несут лишь один интрон, а один ген и вовсе имеет безинтронное строение. Как же так получилось? Можно предположить, что первичный актиновый ген был прерывающимся и все современные актиновые гены произошли от него, попутно теряя свои интроны. Причем в каждой эволюционной ветви были потеряны различные интроны. Возможно, некоторые из интронов были полностью утеряны в эволюции. Таким образом, первичный ген мог содержать 20 интронов или больше. В качестве альтернативной гипотезы можно предположить, что процесс вставки интронов происходил независимо в различных эволюционных линиях. Взаимосвязь между расположением интронов, обнаруживаемая у представителей различных видов, может быть использована для построения эволюционного дерева гена.

Соотношение между экзонами и доменами в белках непостоянно. В некоторых случаях есть четкая взаимосвязь экзон-домен; в других какая-либо связь между доменами и экзонами отсутствует. Возможно, удаление интронов привело к объединению соседних экзонов. В этом случае интрон должен был быть удален очень точно, не затрагивая целостность кодирующего региона. Альтернативой предложенному объяснению может быть возникновение некоторых новых интронов путем вставки в сцепленные домены. Если учесть еще и вариабельность локализации интронов (как мы видели в примере с генами актинов), можно сделать вывод, что расположение интронов может меняться в процессе эволюции.

Возможность сопоставления некоторых экзонов с белковыми доменами и существование гомологичных экзонов в различных генах не оставляют сомнений в том, что дупликации и перемещения экзонов сыграли важную роль в процессе эволюции. Возможно, количество предковых экзонов (из которых в результате дупликаций, вариаций и рекомбинаций произошли все белки), было сравнительно небольшим (несколько тысяч или десятков тысяч). Если воспринимать экзоны как строительный материал эволюции, то приходится принять гипотезу о том, что современные гены возникли по модели «ранних интронов».



РИС. 3.23 Гены актинов сильно варьируют по структуре. Места расположения интронов обозначены красным

3.11 Вся ли генетическая информация содержится в ДНК?

Ключевые понятия:

- В последнее время определение термина «ген» существенно изменилось. Утверждение «один ген — один белок» превратилось в «один белок — один ген».
- Позиционная информация также важна при развитии организма.

Концепция гена особенно бурно развивалась последние несколько лет. Какое же из определений лучше всего подходит понятию ген? Мы больше не можем говорить, что ген — это уникальная последовательность ДНК, кодирующая определенный белок. В случаях, когда участок ДНК отвечает за производство одного белка, современная терминология рассматривает ген как полную последовательность ДНК — от первой точки, представленной в мРНК, до последней точки, соответствующей ее концу, включая экзоны интроны и все остальное.

Если же последовательности, кодирующие разные белки, перекрываются или имеют альтернативные формы сплайсинга, мы можем перевернуть обычное определение и, вместо того, чтобы сказать «один ген — один полипептид», определить их взаимоотношение как «один полипептид — один ген». Таким образом, мы обозначим, что ген — это последовательность (включающая как экзоны, так и интроны), ответственная за производство полипептида, но с оглядкой на то, что часть той же самой последовательности может также принадлежать и *другому* гену. Это позволяет ввести термин «перекрывающегося» или «альтернативного» гена.

Теперь видно, насколько далеко мы отошли от начальной гипотезы «один ген — один фермент». Раньше главным вопросом была сама природа гена. Когда же люди установили, что гены кодируют белки — парадигма зафиксировалась в виде концепции, что функция каждой генетической единицы — это синтез определенного белка.

Это представление отражает центральную догму молекулярной биологии: последовательность ДНК осуществляет свою функцию как за счет прямого кодирования определенного белка, так и за счет влияния прилегающих регионов на собственно кодирующие регионы. Как далеко заведет нас эта парадигма в попытке объяснить простейшие взаимосвязи между генами и белками?

Развитие многоклеточных организмов основано на использовании множества генов для производства различных клеточных фенотипов в каждой ткани. Экспрессия генов регулируется сетью молекулярных механизмов (сигнальных каскадов). Экспрессия первой группы генов, происходящая в самом начале развития эмбриона, приводит к экспрессии генов, вовлеченных в последующие стадии развития, которые, в свою очередь, влияют на гены следующей стадии, и так до тех пор, пока все ткани сформировавшегося организма не начнут работать. Точная молекулярная природа этой регуляторной сети пока неизвестна, однако мы предполагаем, что она базируется на работе генов, кодирующих продукты (белки и РНК), способные оказывать воздействие на другие гены.

Несмотря на превалирующую гипотезу о том, что программа развития организмов осуществляется по каскадному механизму, можно задаться вопросом, достаточно ли она обоснована. Особое внимание здесь надо уделить природе и роли **позиционной информации**. Мы знаем, что части оплодотворенного яйца неравнозначны. Одной из особенностей, влияющих на развитие различных тканей из различных областей яйцеклетки, является именно пространственное расположение неких информационных единиц (предположительно, макромолекул) в пределах клетки.

Мы до конца не понимаем, как формируются изначальные регионы эмбриона. Однако мы можем предположить, что существование позиционной информации в яйцеклетке определяет экспрессию разных генов в клетках, оказавшихся в разных регионах эмбриона. Это, в свою очередь, ведет к развитию взрослого организма, который даст начало новой яйцеклетке со свойственной ей позиционной информацией.

Существование позиционной информации наталкивает нас на мысль: а может ли часть информации, необходимой для развития организма, существовать в форме, отличной от последовательности нуклеотидов ДНК (даже несмотря на необходимость экспрессии определенных участков ДНК для сохранения позиционной информации). В еще более общем виде мы можем спросить себя примерно так: когда мы знаем всю последовательность ДНК, представляющую геном некоторого организма, и умеем интерпретировать полученную информацию в виде белков и регуляторных регионов, можем ли мы (хотя бы теоретически) построить организм (или даже одну живую клетку) путем контролируемой экспрессии нужных генов?

3.12 Резюме

Все типы эукариотических геномов содержат прерывающиеся гены. Доля прерывающихся генов еще невелика у дрожжей и увеличивается у низших эукариот, а у высших эукариот лишь небольшое число генов имеет безынтронную структуру.

Интроны обнаружены во всех классах эукариотических генов. Экзон-интронная структура каждого прерывающегося гена одинакова во всех тканях организма: экзоны соединяются в РНК в том же порядке, в каком они расположены в ДНК, а интроны обычно не несут белок-кодирующей функции. Интроны удаляются из РНК в результате сплайсинга. Некоторые гены используют альтернативные варианты сплайсинга (когда в одном из вариантов некоторая последовательность воспринимается как интрон и удаляется, а вдругом — остается в последовательности в виде экзона).

При межвидовом сравнении гомологичных генов видно, что места расположения интронов часто консервативны. Последовательности интронов высоковариабельны и могут вовсе не обнаруживать сходства даже для случаев, когда последовательности экзонов гомологичны. Привязанность интронов к определенным местам в гене позволяет выявлять гомологичные гены у представителей различных видов.

Размер гена в первую очередь зависит от длины его интронов. Длина интронов увеличивается от низших к высшим эукариотам, что ведет к параллельному увеличению длины генов. У млекопитающих размер генов варьирует в пределах 1-100 тпн. Однако встречаются и более протяженные гены: наиболее известным случаем длинных генов является ген дистрофина, размер которого достигает 2000 тпн.

Некоторые гены гомологичны лишь по некоторым из своих экзонов, и предполагается, что эволюция таких генов шла путем перекомбинирования экзонов, представляющих индивидуальные модули белков. Одинаковые модули могли быть вставлены во множество различных белков. Идея о том, что гены собирались путем комбинирования экзонов, наводит на мысль, что интроны были еще у древнейших организмов. Сопоставление структур некоторых современных генов указывает на существование механизма утери интронов (из структуры предкового гена), причем в разных эволюционных линиях терялись разные интроны.

Литература

3.2 Прерывистый ген состоит из экзонов и интронов

Обзор

Breathnach, R. and Chambon, P. (1981). Organization and expression of eukaryotic split genes coding for proteins. *Annu. Rev. Biochem.* 50, 349–383.

Faustino, N. A. and Cooper, T. A. (2003). Pre-mRNA splicing and human disease. *Genes Dev.* 17, 419–437.

3.3 Эндонуклеазы рестрикции – это один из основных инструментов картирования ДНК

Обзор

Nathans, D. and Smith, H. O. (1975). Restriction endonucleases in the analysis and restructuring of DNA molecules. *Annu. Rev. Biochem.* 44, 273–293.

Wu, R. (1978). DNA sequence analysis. *Annu. Rev. Biochem.* 47, 607–734.

Исследование

Danna, K. J., Sack, G. H., and Nathans, D. (1973). Studies of SV40 DNA VII A cleavage map of the SV40 genome. *J. Mol. Biol.* 78, 363–376.

3.4 Организация прерывистых генов консервативна

Исследование

Berget, S. M., Moore, C., and Sharp, P. (1977). Spliced segments at the 5′ terminus of adenovirus 2 late mRNA. *Proc. Natl. Acad. Sci. USA* 74, 3171–3175.

Chow, L. T., Gelinas, R. E., Broker, T. R., and Roberts, R. J. (1977). An amazing sequence arrangement at the 5′ ends of adenovirus 2 mRNA. *Cell* 12, 1–8.

Glover, D. M. and Hogness, D. S. (1977). A novel arrangement of the 8S and 28S sequences in a repeating unit of D. melanogaster rDNA. *Cell* 10, 167–176.

Jeffreys, A. J. and Flavell, R. A. (1977). The rabbit b-globin gene contains a large insert in the coding sequence. *Cell* 12, 1097–1108.

Wenskink, P. et al. (1974). A system for mapping DNA sequences in the chromosomes of *D. melanogaster*. *Cell* 3, 315–325.

3.9 Некоторые экзоны напоминают последовательности, кодирующие целые белки

Обзор

Blake, C. C. (1985). Exons and the evolution of proteins. *Int. Rev. Cytol.* 93, 149–185.

Структура генома

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

4.1 Введение

Наиболее важной характеристикой генома является число составляющих его генов. Говоря о генах в геноме, можно выделить четыре уровня реализации генетической информации:

- **Геном** — это полный набор генов организма. Геном представляет собой полную последовательность ДНК организма (хотя на практике невозможно точно локализовать каждый ген, основываясь только на последовательности нуклеотидов).
- **Транскриптом** — это полный набор молекул РНК, синтезируемых в определенных условиях. Транскриптом может относиться как к однотипным клеткам, так и к более разнородной клеточной системе вплоть до целого организма. Поскольку некоторые гены кодируют множество мРНК, исчерпывающий транскриптом организма должен быть несколько больше числа генов в геноме. Транскриптом включает как мРНК, так и не кодирующие белок РНК (иными словами, все, что транскрибируется).
- **Протеом** — это полный набор белков (организма или отдельных тканей и клеток). Он должен соответствовать мРНК в транскриптоме (хотя для конкретных условий протеом и транскриптом могут отличаться за счет колебания относительной представленности или стабильности мРНК и белков). Протеом можно определить как исчерпывающий набор белков, кодируемых геномом, или как набор белков, производимых в каждой отдельной клетке или ткани.
- Белки могут функционировать как независимо, так и в составе белковых комплексов. Если бы мы точно знали про все белок-белковые взаимодействия, мы смогли бы определить общее количество независимых белковых комплексов.

Количество генов в геноме можно определить по числу открытых рамок считывания. Однако локализация генов по этому признаку является довольно сложной задачей, поскольку прерывистые гены эукариот состоят из фрагментированных (интронами) рамок считывания. Более того, при таком подходе мы чаще всего не получаем представления о функциях предполагаемых белковых продуктов и не имеем доказательств того, что выявленные гены вообще экспрессируются. Поэтому данный подход непригоден для определения *потенциала* генома. Тем не менее есть серьезные основания полагать, что любая консервативная рамка считывания должна экспрессироваться.

Другой подход к определению количества генов в геноме осуществляется в терминах транскриптома (путем прямого определения всех мРНК) или протеома (путем определения белков). Такие подходы дают гораздо большую уверенность, что мы имеем дело с генами, которые на самом деле экспрессируются. Исследование транскриптома или протеома позволяет нам выяснить, сколько генов экспрессируется в определенном типе клеток или тканей, какие есть отличия в относительном уровне экспрессии и сколько генов уникально экспрессируются в данной клетке по сравнению с любой другой клеткой.

Чтобы выяснить, насколько конкретный ген важен для жизнедеятельности, необходимо посмотреть, что происходит в случае его нуль-мутации. Если мутация летальна или у организма присутствует видимый дефект, можно сделать вывод, что ген жизненно важен или, по крайней мере, обеспечивает селективное преимущество. Однако многие гены могут быть удалены без видимого эффекта на фенотип. Действительно ли такие гены неважны или же отсутствие гена проявится в других обстоятельствах или в течение более длительного времени?

4.2 Геномы можно картировать при помощи методов классической генетики, расщепления рестриктазами или путем прямого секвенирования ДНК

Характеристику генома важно начать с составления его генетической карты. Картирование генов и геномов можно рассматривать на нескольких уровнях:

- Классическая генетическая карта (или карта сцепления) определяет расстояние между мутациями с позиции частот рекомбинации. Она ограничена тем, что опирается на возникновение мутаций, проявляющихся в фенотипе. Поскольку частоты рекомбинации не всегда соответствуют физическому расстоянию между участками, такая карта не очень точно отражает длину ДНК между двумя точками.
- Карта сцепления может быть также построена не на фенотипически проявляющихся мутациях, а на частоте рекомбинаций между отдельными участками ДНК (сайтами рестрикции). Эти участки обладают вариабельностью, порождающей различия в их узнавании некоторыми ферментами (рестриктазами). Поскольку такой полиморфизм довольно распространен, подобные карты могут быть сделаны для любого организма (вне зависимости от возможности получения мутантов). Однако такая карта имеет те же недостатки, что и любая другая карта сцепления, — относительные расстояния на ней основаны не на физической длине, а на рекомбинации.
- Карта рестрикции создается путем разрезания ДНК на фрагменты с помощью эндонуклеаз рестрикции (рестриктаз) с последующим определением длины получаемых фрагментов. В данном случае расстояния выражаются через длину фрагментов ДНК, и, таким образом, картирование дает физическую карту генетического материала. На самом деле, рестрикционная карта сама по себе не позволяет локализовать интересующие участки генома. Ее ассоциация с генетической картой происходит за счет мутаций в сайтах, узнаваемых эндонуклеазами рестрикции. При помощи рестрикционной карты могут быть распознаны значительные изменения в геноме (делеции и инсер-

ции), поскольку они влияют на размер или количество рестрикционных фрагментов. Однако точечные мутации можно выявить лишь при их расположении непосредственно в сайте, узнаваемом ферментом.

- Самая надежная генетическая карта основана на определении последовательности нуклеотидов в ДНК (секвенировании ДНК). Зная полную последовательность ДНК, мы можем локализовать гены и определить расстояния между ними. Путем сравнения (выравнивания) нескольких схожих последовательностей можно сделать вывод об их способности кодировать белок. Простейшим допущением здесь является то, что благодаря естественному отбору мутации (как правило, негативно влияющие на свойства белков) накапливаются в кодирующей последовательности гораздо медленнее, чем за пределами кодирующих регионов ДНК. Перевернув этот аргумент, можно сказать, что более консервативная последовательность, скорее, является кодирующей.

Сравнивая между собой последовательности ДНК аллеля дикого типа и мутантного аллеля, можно точно локализовать мутационный сайт и установить природу мутации. Таким образом, можно установить взаимосвязь между генетической картой (полностью основанной на мутационных сайтах) и физической картой (основанной на нуклеотидной последовательности ДНК или собственно являющейся последовательностью ДНК).

Перечисленные методы используют как для локализации отдельных кодирующих регионов, так и для картирования целых геномов (хотя, конечно, масштабы таких карт сильно различаются). Независимо от используемого метода картирования, принцип всех подходов заключается в том, чтобы получить некий набор перекрывающихся регионов ДНК с тем, чтобы потом соединить их в протяженную карту. Необходимым условием создания генетических карт является то, что каждый сегмент карты должен быть связан с соседними сегментами строго специфичными участками перекрывания по ДНК. В этом случае можно быть уверенным, что ни один из сегментов не утерян. Этот принцип применяется как на уровне создания крупномасштабных (хромосомных) карт, так и при сборке коротких фрагментов внутри более крупных регионов.

4.3 Геномы отличаются огромным разнообразием

Ключевые понятия:

- Полиморфизм можно обнаружить на уровне фенотипа (когда варианты последовательности гена влияют на функцию этого гена), на уровне рестрикционных фрагментов (когда полиморфизм последовательности ДНК затрагивает мишень рестриктазы), и на уровне собственно последовательности нуклеотидов (когда полиморфизм последовательностей выявляют путем прямого секвенирования ДНК).
- Аллели генов демонстрируют значительный полиморфизм на уровне последовательности, однако многие вариации не влияют на функцию.

Классическая менделевская генетика разделяла встречающиеся в популяции варианты гена на аллели дикого типа (обычные, часто встречающиеся в популяции) и мутантные аллели (новые варианты, редко встречающиеся). Позже исследователи обнаружили, что для многих генов в популяции присутствует несколько вариантов примерно в равной пропорции, ни один из которых нельзя с уверенностью назвать диким типом (предковым аллелем). Тогда была сформулирована гипотеза так называемых множественных аллелей, каждый из которых оказывает собственное влияние на фенотип и является диким типом. Во многих случаях выделение одного из вариантов гена в качестве дикого типа оказывается крайне сложной задачей.

Существование множественных аллелей одного локуса называют генетическим **полиморфизмом**. Каждый генетический локус, несколько аллелей которого представлены в популяции, по определению полиморфен. В настоящее время принято относить аллель к «нормальным» (не мутантным), если частота его встречаемости в популяции превышает 1%.

В чем же заключается генетическая суть аллельного полиморфизма? Множественные аллели содержат различия (возникшие в эволюции в результате мутаций), способные влиять на функцию белка и, таким образом, изменять фенотип. Если мы сравним рестрикционные карты нуклеотидных последовательностей множественных аллелей, они тоже будут полиморфны (карта каждого аллеля будет отличаться от остальных).

Как уже было сказано, полиморфным может быть и вариант дикого типа (если полиморфизм не сказывается на фенотипе). В настоящее время аллели дикого типа можно выявить по отличиям в нуклеотидной последовательности, даже не влияющим на кодируемую функцию (и не приводящим к изменению фенотипа). Популяции могут быть весьма полиморфны на уровне генотипа. В данном локусе могут встречаться различные варианты нуклеотидной последовательности. Некоторые из них влияют на фенотип, а некоторые «скрыты», поскольку не оказывают видимого эффекта.

Почти в каждом генетическом локусе можно обнаружить набор нуклеотидных замен (вариантов) следующих типов: 1) изменяющих нуклеотидную последовательность ДНК, но не влияющих на структуру белка; 2) изменяющих последовательность белка, не меняя его функции; и 3) являющихся причиной возникновения белков, демонстрирующих различные свойства (или приводящих к формированию мутантных, нефункциональных белков).

Если сравниваемые аллели отличаются одним нуклеотидом, то такие замены называют **однонуклеотидным полиморфизмом (ОНП)**. Для генома человека в среднем на 1330 оснований приходится 1 полиморфный нуклеотид. Каждый человек уникален по своему набору ОНП. Существует множество способов детекции ОНП, начиная с прямого сравнения нуклеотидных последовательностей и заканчивая масс-спектрометрией или биохимическими подходами, выявляющими отличия на определенном участке ДНК.

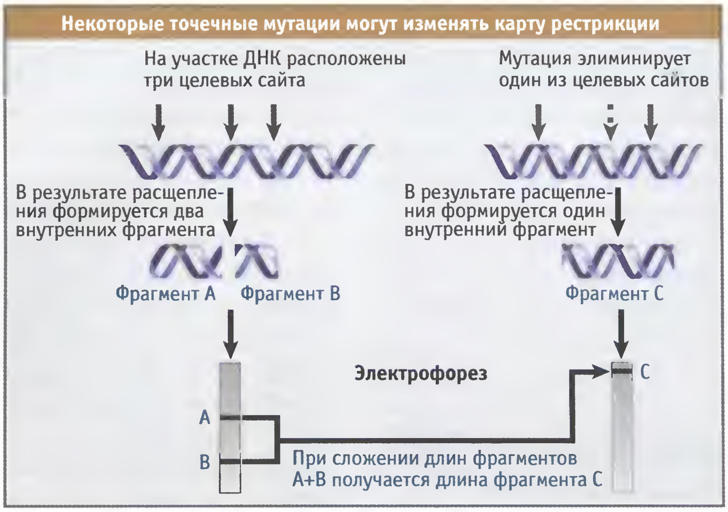


РИС. 4.1. Точечную мутацию в сайте рестрикции можно обнаружить по разнице в длине получаемых фрагментов

Одной из задач генетического картирования является получение списка наиболее распространенных вариантов. Частота встречаемости ОНП в геноме позволяет предсказать, что в целом (суммируя геномы всех живущих людей) в популяции людей должно быть не менее 10 миллионов ОНП, встречающихся с частотой >1% (в поплуляции). На сегодняшний день известно уже более миллиона из них.

Некоторые полиморфизмы можно детектировать, сравнивая рестрикционные карты различных индивидуумов. Критерием при этом является изменение набора фрагментов ДНК, получаемых в результате расщепления рестриктазой. Когда мишень для фермента (сайт рестрикции) есть в геноме одного человека и отсутствует в геноме другого, расщепление первого генома приведет к образованию двух фрагментов, соответствующих лишь одному фрагменту второго генома (**РИС. 4.1**).

Карта рестрикции не зависит от функции гена, поэтому рестрикционный полиморфизм может быть обнаружен *вне зависимости от того, оказывает ли последовательность влияние на фенотип*. Лишь немногие расположенные в сайтах рестрикции полиморфизмы могут влиять на фенотип. Большинство из них приводят



РИС. 4.2. Полиморфизм сайтов рестрикции наследуется в соответствии с законами Менделя. Четыре аллеля рестрикционного маркера встречаются во всех возможных парных комбинациях и разделяются независимо в каждом поколении. Фотография любезно предоставлена Ray White, Ernest Gallo Clinic and Research Center, University of California, San Francisco

к изменениям, не оказывающим влияния на производство белка (например, потому, что они лежат в некодирующих регионах генома).

Различия в картах рестрикции двух индивидуумов называют **полиморфизмом длин рестрикционных фрагментов (ПДРФ)**. Проще говоря, ПДРФ — это следствие ОНП, расположенных в сайте узнавания эндонуклеазы рестрикции. ПДРФ можно использовать в качестве генетического маркера (так же, как и любой другой маркер). Вместо анализа особенностей фенотипа мы, с помощью карты рестрикции, напрямую оцениваем генотип. На РИС. 4.2 показана генеалогия рестрикционного полиморфизма, отслеженная в трех поколениях. Она демонстрирует менделевское распределение на уровне маркерных фрагментов ДНК.

4.4 ПДРФ и ОНП можно использовать для генетического картирования

Ключевые понятия:

- ПДРФ и ОНП могут дать основу для составления генетических карт, и, кроме того, они подходят для определения взаимосвязи между родителями и потомством.

Можно измерить частоту рекомбинации между рестрикционным и фенотипическим маркерами (РИС. 4.3). Следовательно, полученная таким образом генетическая карта будет содержать как генотипические, так и фенотипические маркеры.

Рестрикционные маркеры не ограничены изменениями генома, влияющими на фенотип, и поэтому они являются очень мощным инструментом для изучения генетического локуса на молекулярном уровне. Типичной проблемой современной генетики является существование мутаций, оказывающих эффект на фенотип, но для которых неизвестен соответствующий ген или белок. Множество опасных и даже смертельных заболеваний человека попадают в эту категорию. Так, например, муковисцидоз демонстрирует менделевский тип наследования, но молекулярная природа мутации была неизвестна до тех пор, пока не был охарактеризован соответствующий ген.

Если рестрикционный полиморфизм распределен по геному случайно (и более или менее равномерно), некоторые полиморфные участки могут располагаться рядом с геном, определяющим заболевание. В этом случае можно попробовать локализовать такие рестрикционные маркеры благодаря тому что они ассоциированы с мутантным фенотипом. Если сравнить рестрикционные карты ДНК пациентов, страдающих от некоторого генетического заболевания, с ДНК здоровых людей, можно обнаружить, что у пациентов определенные сайты рестрикции либо всегда есть, либо их всегда нет.

Гипотетический пример показан на РИС. 4.4. Приведенная ситуация отражает полную взаимосвязь между рестрикционным маркером и фенотипом. Следовательно, рестрикционный маркер расположен так близко к мутантному аллелю, что они никогда не разделяются в результате рекомбинации.

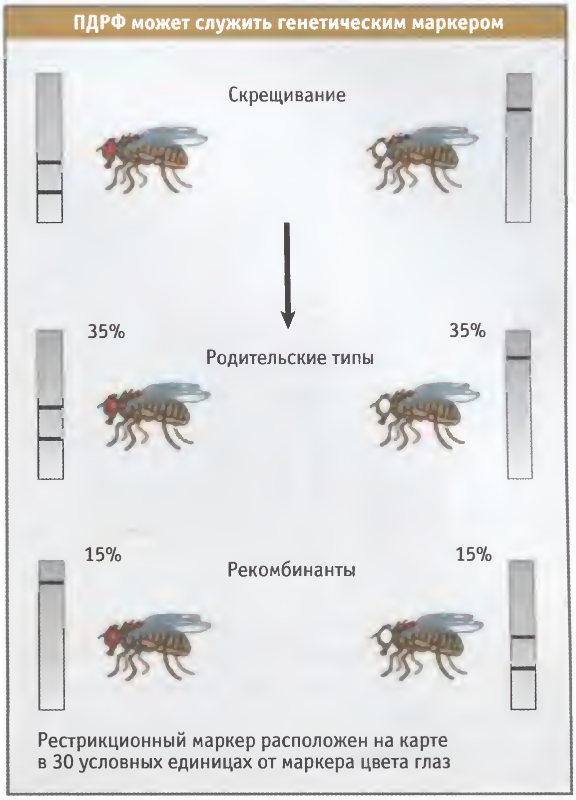


РИС. 4.3. Полимофизм сайтов рестрикции можно использовать в качестве генетического маркера при измерении расстояния от фенотипического маркера (например, маркера цвета глаз) до места рекомбинации. На рисунке в упрощенном виде показана такая ситуация на примере полос ДНК, соответствующих аллелю одного гена в диплоидном организме

Локализация такого маркера имеет 2 важных последствия:

- Знание рестрикционного маркера генетического заболевания может стать основой процедуры обнаружения данного заболевания. Некоторые хорошо изученные с медицинской точки зрения заболевания, но для которых пока неизвестна молекулярно-генетическая основа, очень сложно диагностировать. Если рестрикционный маркер надежно связан с фенотипом, его наличие у пациента может быть использовано при диагностике заболевания.
- Знание рестрикционного маркера генетического заболевания может привести к обнаружению собственно гена, отвечающего за измененный фенотип. Рестрикционный маркер должен находиться на генетической карте в относительной близости от гена, чтобы между ними редко (или никогда) не оказывался сайт рекомбинации. «Относительно близко» в терминах генетики может быть довольно значительным расстоянием в терминах пар оснований ДНК;



РИС. 4.4. Если маркер рестрикции ассоциирован с фенотипическим признаком, сайт рестрикции должен располагаться вблизи гена, ответственного за данный фенотип. Мутации, меняющие маркеры, характерные для пациентов, могут указывать на локализацию гена заболевания

тем не менее выявленный рестрикционный маркер является стартовой точкой, от которой мы можем продвигаться по ДНК к гену.

Высокая частота ОНП в человеческом геноме делает их удобным инструментом генетического картирования. Для полутора миллионов уже известных ОНП частота встречаемости в среднем равна одному ОНП на 1–2 тпн. Такая частота генетических маркеров позволяет быстро локализовывать новые гены заболеваний путем выявления ассоциированных с заболеванием ОНП.

Для тех же целей некоторое время назад использовали и ПДРФ-картирование. Поскольку ПДРФ жестко ассоциирован с некоторой группой сцепления, он может быть нанесен на генетическую карту. ПДРФ-картирование человека и мыши привело к построению карт геномов этих организмов. Любой ранее неизвестный сайт может быть протестирован на сцепление и быстро нанесен на карту. Число регионов ПДРФ гораздо меньше числа ОНП, поэтому разрешение карты ПДРФ ниже.

Высокая частота полиморфных регионов генома приводит к тому, что каждый индивидуум уникален по набору своих ОНП или ПДРФ. Частное сочетание сайтов, расположенных в определенном регионе ДНК, называют **гаплотипом** (генотипом в миниатюре). Гаплотип изначально был представлен в виде концепции для описания генетического состава локуса главного комплекса гистосовместимости — региона, кодирующего важные белки иммунной системы (см. гл. 23 «Иммунное разнообразие»). В настоящее время эта концепция получила распространение и используется при описании определенной комбинации аллелей (или любых других генетических маркеров, например сайтов рестрикции), находящихся в определенном месте генома. С использованием ОНП была получена детальная карта гаплотипов человеческого генома. Это значительно облегчает картирование генов, связанных с возникновением заболеваний.

Существование ПДРФ является удобным инструментом для определения степени родства. Так, в случаях, когда отцовство вызывает сомнения, сравнение карт ПДРФ (на некотором участке хромосомы) для детей и предполагаемых родителей позволяет достоверно установить биологическое родство. Использование рестрикционного анализа геномной ДНК для идентификации личности получило название **генетического отпечатка пальца** (по аналогии с отпечатками пальцев, **DNA fingerprinting**). Анализ высоковариабельных последовательностей так называемых «мини-сателлитов» был использован и при картировании генома человека (см. разд. 6.14 «Мини-сателлиты можно использовать при генетическом картировании»).

4.5 Почему геномы такие большие?

Ключевые понятия:

- Четкая корреляция между размером генома и генетической сложностью отсутствует.
- Минимальный размер генома возрастает пропорционально увеличению сложности организма.
- Размеры геномов могут сильно варьировать даже в пределах одного таксона.

Общее количество ДНК в гаплоидном гене является характерной чертой каждого вида и называется

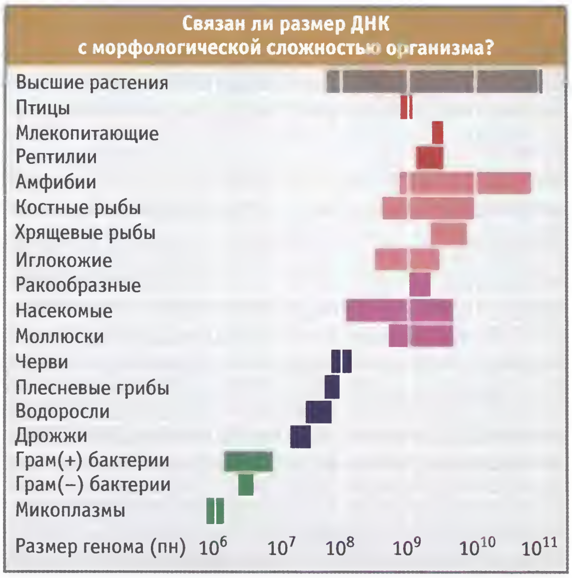


РИС. 4.5. Размер ДНК гаплоидного генома увеличивается по мере морфологического усложнения низших эукариот, однако в пределах групп высших эукариот он может варьировать в широком диапазоне. Закрашенные участки отражают разброс длин геномов в пределах каждой группы

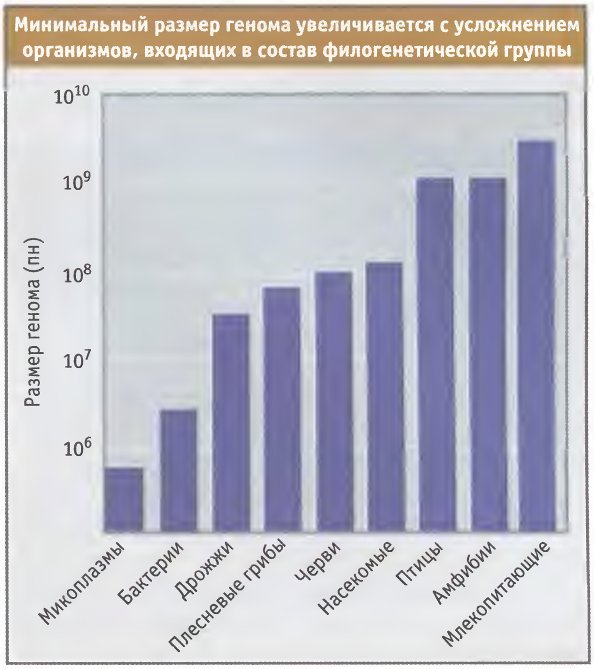


РИС. 4.6. Минимальный размер генома в пределах филогенетической группы увеличивается от прокариот к млекопитающим

величиной С. Величина С колеблется в широких пределах, от менее чем 10^6 пн у микоплазмы до свыше 10^{11} пн у некоторых растений и амфибий.

На **РИС. 4.5** приведен диапазон значений величины С для организмов, находящихся на разных ступенях эволюционного развития. С усложнением организации в каждой группе организмов наблюдается увеличение минимального (для данной группы организмов) размера генома. Одновременно с увеличением количества ДНК увеличивается и разброс значений величины С в пределах одного таксона (у высших эукариот можно наблюдать большие различия даже для близкородственных видов).

Гистограмма, отображающая *минимальное* количество ДНК, обнаруживаемое в пределах одного таксона (**РИС. 4.6**), позволяет предположить, что для прокариот и низших эукариот необходимым условием увеличения сложности организма является увеличение размера его генома.

Геном одного из самых маленьких прокариот — микоплазмы — всего в три раза больше крупного генома бактериофага. Размер геномов бактерий начинается со значения 2×10^6 пн. Геномы одноклеточных эукариот (чей жизненный цикл во многом напоминает прокариотический) превосходят по размеру геномы бактерий, но меньше геномов многоклеточных. Принадлежность к эукариотам как таковая не означает значительного увеличения размеров генома. Например, размер генома дрожжей составляет около $1,3 \times 10^7$ пн, что лишь в два раза превышает средний размер бактериального генома.

Дальнейшее двукратное увеличение размеров генома оказывается достаточным для того, чтобы миксомицеты *Dictyostelium discoideum* могли существовать как в виде одноклеточных, так и в виде многоклеточных организмов. Для возникновения первых по-настоящему многоклеточных организмов необходимо дальнейшее усложнение генома. Так, например, у нематоды *Caenorhabditis elegans* количество ДНК достигает 8×10^7 пн.

Также можно отметить зависимость увеличения размеров генома от сложности организации в ряду наиболее изучаемых (модельных) организмов, представленную на **РИС. 4.7**. Так, последовательное увеличение размеров генома необходимо для появления насекомых, птиц, амфибий и млекопитающих. Тем не менее начиная с млекопитающих четкая взаимосвязь между размером генома и морфологической сложностью теряется.

Размеры геномов, которые полезно знать

Таксон	Вид	Геном (пн)
Водоросли	*Pyrenomas salina*	$6,6 \times 10^5$
Микоплазмы	*M. pneumoniae*	$1,0 \times 10^6$
Бактерии	*E. coli*	$4,2 \times 10^6$
Дрожжи	*S. cerevisiae*	$1,3 \times 10^7$
Миксомицеты	*D. discoideum*	$5,4 \times 10^7$
Нематоды	*C. elegans*	$8,0 \times 10^7$
Насекомые	*D. melanogaster*	$1,4 \times 10^8$
Птицы	*G. domesticus*	$1,2 \times 10^9$
Амфибии	*X. laevis*	$3,1 \times 10^9$
Млекопитающие	*H. sapiens*	$3,3 \times 10^9$

РИС. 4.7. Размер геномов некоторых модельных организмов

Мы знаем, что гены высших эукариот по размеру значительно больше, чем последовательности, собственно кодирующие белок, поскольку экзоны (кодирующие регионы) могут составлять лишь малую часть от общей длины гена. Это отчасти объясняет, почему общее количество ДНК больше, чем необходимо для существования рамок считывания всех белков организма. Большие области прерывистых генов могут быть не связаны с функцией кодирования белков (являясь интронами). Дополнительно существуют и значительные участки ДНК, расположенные между генами. Таким образом, общий размер генома не дает представления о количестве входящих в него генов.

Слабая корреляция между размером генома и сложностью генетического устройства получила название **парадокса величины С**. Есть очень любопытные примеры вариации размеров генома. Например, у лягушки *Xenopus* и человека геномы практически одинакового размера. При этом мы считаем, что в терминах генетического развития человек устроен более сложно! Существуют таксоны, внутри которых организмы демонстрируют значительные отличия в количестве ДНК, не обнаруживая при этом отличий в сложности организации (рис. 4.5). (Это особенно заметно у насекомых, амфибий и растений, и практически не проявляется у птиц, рептилий и млекопитающих, у которых геномы в пределах группы варьируют не более чем в 2 раза.) У сверчка ДНК в 11 раз длиннее, чем у плодовой мушки. У амфибий размер генома варьирует в диапазоне значений от 10^9 пн до 10^{11} пн. Вряд ли амфибии с максимальным и минимальным размерами геномов сильно отличаются по числу генов. Еще не выяснено, почему естественный отбор допускает существование таких вариаций и какие это имеет последствия для эволюции.

4.6 Эукариотические геномы состоят из уникальных и повторяющихся последовательностей

Ключевые понятия:

- Кинетика реассоциации молекулы ДНК после денатурации позволяет разделить последовательности по частоте их встречаемости в геноме.
- Подавляющее большинство генов — это уникальные для генома последовательности.
- В пределах одного таксона более крупные геномы содержат не большее число генов, а большее число повторяющихся участков ДНК.
- Существенная часть повторяющихся участков ДНК представляет собой транспозоны.

ДНК эукариотического генома можно охарактеризовать с помощью кинетики реассоциации денатурированной ДНК. Этот метод широко использовался до тех пор, пока не появилась возможность эффективного секвенирования ДНК.

Кинетика реассоциации выделяет два основных типа геномных последовательностей:

- **Уникальные последовательности,** встречающиеся один раз на гаплоидный геном.
- **Повторяющиеся последовательности** — регионы, представленные в геноме более чем одной копией.

Повторяющиеся последовательности ДНК в свою очередь принято разделять на два основных типа:

- Умеренно повторяющиеся последовательности представлены регионами различной длины, встречающимися в геноме в количестве 10–1000 копий. Эти последовательности разбросаны по геному и отвечают за формирование вторичной структуры пре-мРНК (когда инвертированные повторы интронов соединяются попарно, формируя дуплексы).
- Высокоповторяющиеся последовательности ДНК представлены очень короткими участками (обычно менее 100 пн), которые многие тысячи раз встречаются в геноме и часто организованы в виде тандемных повторов (см. разд. 6.11 «Сателлитная ДНК часто расположена в гетерохроматине»). Высокоповторяющиеся последовательности не кодируют белки.

Уникальные последовательности ДНК могут составлять от нескольких до 100% от всего генома. На РИС. 4.8 изображена организация геномов некоторых модельных организмов. Прокариоты содержат практически только уникальную ДНК. У низших эукариот уникальные последовательности составляют подавляющую часть генома (только <20% последовательностей попадает в категорию более или менее умеренно повторяющихся регионов). В клетках животных умеренно- и высокоповторяющиеся участки занимают до половины ДНК. У растений

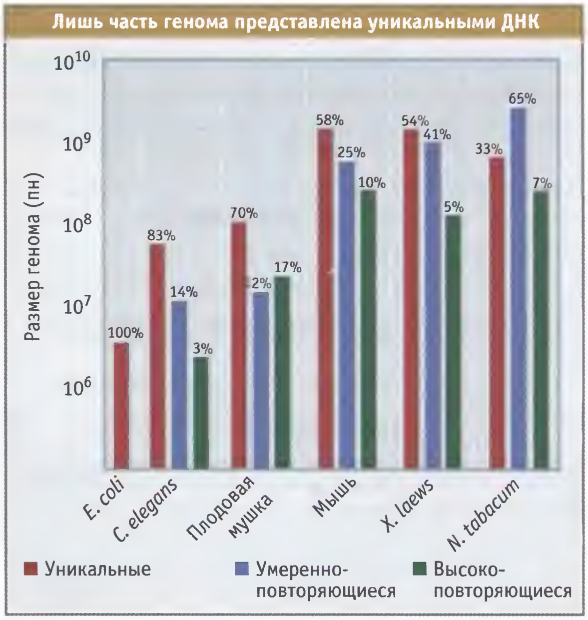


РИС. 4.8. В эукариотических геномах соотношение различных компонентов последовательности сильно варьирует. Общий размер уникальных последовательностей ДНК возрастает с увеличением размера генома, достигая плато на уровне ~2×10^9 пн

и амфибий они могут составлять до 80% генома, и, таким образом, уникальная ДНК у этих организмов находится в меньшинстве.

Заметная часть умеренно повторяющейся ДНК состоит из **транспозонов** — коротких последовательностей (около 1 тпн), способных перемещаться в новые места генома и/или производить дополнительные копии самих себя (см. гл. 21 «Транспозоны» и гл. 22 «Ретровирусы и ретропозоны»). У некоторых высших эукариот они могут занимать более половины генома (см. разд. 5.5 «В геноме человека меньше генов, чем кажется»).

Транспозоны иногда рассматривают в контексте концепции **эгоистичной ДНК** (selfish DNA) в качестве самоподдерживающихся последовательностей, не участвующих в процессе развития организма. Транспозоны могут способствовать появлению геномных перестроек, которые, в свою очередь, могут приводить к некоторым преимуществам данного организма в процессе естественного отбора. К сожалению, до конца непонятно, почему естественный отбор не противодействует такому колоссальному накоплению транспозонов в геномах.

Еще один термин, используемый для объяснения избыточности ДНК, — это *мусорная ДНК* (junk DNA), к которой относят геномные последовательности без известной функции. Вероятнее всего, в геноме существует баланс между возникновением новых последовательностей и удалением ненужных, и та часть ДНК, для которой мы не обнаруживаем функции, может находиться в процессе эволюции или удаления.

Суммарный размер уникальных последовательностей ДНК имеет тенденцию к увеличению с возрастанием размера генома, вплоть до характерного для млекопитающих размера в 3×10^9 пн. Дальнейшее же увеличение размера генома идет за счет возрастания доли повторяющихся компонентов (организмы с размером уникальной ДНК более 2×10^9 пн практически не встречаются). Таким образом, увеличение размера уникального компонента ДНК коррелирует в нашем понимании с относительной сложностью организма. У *E. coli* уникальные элементы занимают $4{,}2 \times 10^6$ пн, у *C. elegans* это значение доходит до $6{,}6 \times 10^7$ пн, у *D. melanogaster* увеличивается уже до 10^8 пн, и наконец, у млекопитающих достигает значений порядка 2×10^9 пн.

Какой же из перечисленных типов ДНК кодирует белки? Кинетика реассоциации ДНК доказывает, что мРНК кодируется уникальными последовательностями генома. Таким образом, количество уникальной ДНК служит лучшим индикатором кодирующего потенциала, чем суммарное количество геномной ДНК (однако более детальный анализ, основанный на известных последовательностях полных геномов, показывает, что множество экзонов содержат участки, родственные другим экзонам (см. разд. 3.5 «Последовательности экзонов консервативны, а интронов — вариабельны»); такие экзоны возникли путем дупликации, в результате которой в одном геноме оказались две изначально идентичные копии, накопившие отличия в ходе эволюции).

4.7 Гены можно локализовать, используя консервативность их экзонов

Ключевые понятия:

- Эволюционно консервативные экзоны можно использовать в качестве основы для определения кодирующих регионов неизученного гена (путем его сравнения с уже исследованными последовательностями других организмов).

Один из стандартных способов определения структуры генов основан на контрасте между консервативностью экзонов и изменчивостью интронов. На участке последовательности, содержащей эволюционно консервативный ген, кодирующие белок регионы должны отвечать двум требованиям:

- Они должны иметь открытую рамку считывания и
- В качестве не обязательного, но желательного пункта они должны иметь гомологов у представителей других видов.

Эти свойства можно использовать для локализации генов и определения их структуры.

Предположим, на основании методов классической генетики мы знаем, что определенный генетический признак расположен в данном регионе хромосомы. Если мы ничего не знаем о структуре продукта гена, как нам локализовать ген на участке, который может достигать, например, 1 Мпн в длину?

Мощный подход, доказавший свою состоятельность в исследовании некоторых генов, имеющих медицинское значение, заключается в анализе относительно коротких фрагментов из региона, обладающего двумя вышеуказанными свойствами консервативного гена. Для начала находят фрагменты, гибридизующиеся с участками геномов других видов, а затем проверяют эти фрагменты на наличие открытых рамок считывания.

Соответствие первому критерию выясняется проведением **зооблота** (zoo blot). Используя короткие фрагменты ДНК изучаемого региона в качестве (радиоактивных) проб, выполняют саузерн-блотинг (Southern blotting) с ДНК разных видов на наличие родственных последовательностей. Если у представителей нескольких видов находятся фрагменты, родственные нашей пробе (обычно соответствующей ДНК человека), то эта проба становится кандидатом на роль экзона в гене.

Фрагменты-кандидаты секвенируют и, если они содержат открытую рамку считывания, используют для выделения прилегающих регионов. Полученный экзон мы можем впоследствии использовать для определения всего гена, для выделения соответствующей кДНК или мРНК и, в итоге, для определения последовательности белка.

Этот подход особенно важен в тех случаях, когда исследуемый ген содержит много больших интронов и в силу этого растянут по геному (как, например, в случае с мышечной дистрофией Дюшенна (МДД)). МДД — это врожденное дегенеративное нарушение мышечной тка-

ни, сцепленное с Х-хромосомой и поражающее одного из 3500 мужчин. Шаги по локализации этого гена приведены на **РИС. 4.9**.

Анализ групп сцепления позволил выявить расположение локуса МДД. Он оказался расположен в локусе Хр21. У пациентов с МДД хромосомные перестановки часто затрагивают именно этот регион Х хромосомы. Путем сравнения способности ДНК-проб Х-хромосом больных индивидов гибридизоваться с ДНК здоровых людей, были получены клоны, соответствующие участку, который претерпел реорганизацию (или был удален в ДНК больных МДД).

Как только были получены последовательности фрагментов, находящихся поблизости от искомого гена, появилась возможность пройти вдоль хромосомы с тем, чтобы выявить находящиеся поблизости гены. «Прогулка» по хромосоме (Genome walking) была использована для создания рестрикционной карты региона по обеим сторонам от начальной пробы (в итоге покрыв регион >100 тпн).



РИС. 4.9. Гены, принимающие участие в развитии миодистрофии Дюшенна, были обнаружены путем хромосомного картирования и исследования участков, несущих делеции, ассоциированные с развитием заболевания

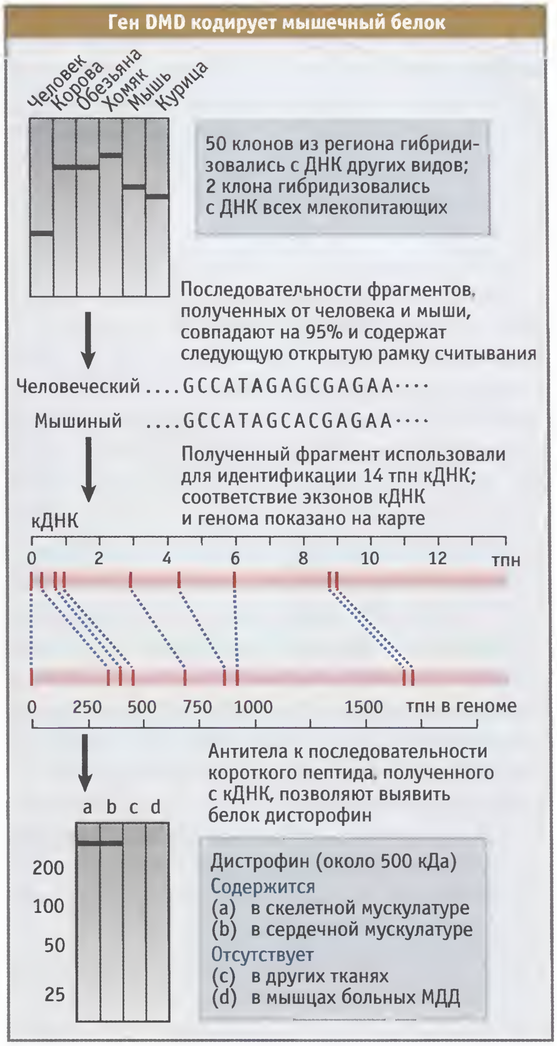


РИС. 4.10. Ген, ответственный за развитие миодистрофии Дюшенна, был охарактеризован с помощью зоо-блотинга, гибридизации кДНК, геномной гибридизации и определения белка

Анализ ДНК, полученной от группы пациентов, позволил выявить крупные делеции в этом регионе, простирающиеся в обоих направлениях. Самой информативной делецией оказалась та, которая целиком лежала в пределах картированного региона. Поскольку она «удаляла» фрагмент генома, важный для нормального функционирования искомого гена, это доказывало, что ген (или, по крайней мере, его часть) лежит именно в этом регионе.

Теперь, локализовав ген, мы должны определить его экзоны и интроны. Зоо-блотинг позволил определить фрагменты, гибридизующиеся с Х-хромосомой мышей и ДНК других млекопитающих. Как показано на **РИС. 4.10**, эти фрагменты были тщательно изучены на наличие открытых рамок считывания и последовательностей, типичных для соединений экзонов и интронов. Фрагменты, соответствовавшие данным критериям, были использованы в качестве проб для определения гомологичных последовательностей в библиотеке кДНК, приготовленной из мРНК мышей.

Оказалось, что соответствующая искомому гену мРНК представляет собой необычно большую молекулу длиной приблизительно 14 000 нуклеотидов. Обратная гибридизация на геном показала, что мРНК закодирована более чем 60 экзонами, которые разбросаны по участку ДНК длинной 2000 тпн. Это означало, что ген МДД — самый крупный из известных генов человека.

Ген кодирует белок дистрофин массой около 500 кДа. Он является компонентом мышечной ткани и представлен в ней в довольно малых количествах. У всех больных людей в локусе этого гена есть делеции, ведущие к отсутствию (или наличию дефектного) дистрофина.

Отличительной чертой мышечной ткани служит наличие еще одного из гигантского белка — титина. Он состоит почти из 27 000 аминокислот. Ген титина содержит наибольшее число экзонов (178) и самый длинный экзон в геноме человека (17 000 пн).

Еще один метод, позволяющий быстро исследовать фрагменты генома на наличие экзонов, называют **улавливанием экзонов** (exon trapping). На РИС. 4.11 показан пример данного подхода. Метод начинается с создания сплайсинг-вектора, содержащего сильный промотор и расположенные за ним два модельных экзона, разделенных интроном. После трансфекции (введения) вектора в клетку его транскрипция приводит к формированию большого количества РНК, содержащей последовательности этих двух экзонов. Внутри интрона расположен специальный рестрикционный сайт, используемый для вставки фрагментов генома из интересующего региона. Если клонированный фрагмент не содержит экзона, то в схеме сплайсинга изменений не происходит, и РНК содержит только те (модельные) последовательности, которые изначально были в векторе. Если же добавленный фрагмент генома содержит целый экзон (фланкированный хотя бы частичными последовательностями интронов), это приводит к распознаванию сайтов сплайсинга по обеим сторонам этого экзона и его вставке в РНК между двумя «модельными» экзонами вектора. Факт вставки можно обнаружить путем обратной транскрипции РНК в кДНК с последующей полимеразной цепной реакции (ПЦР), приводящей к амплификации участка между двумя экзонами вектора. Выявление продукта ПЦР, полученного с клонированного фрагмента генома, означает, что добавленный регион содержит экзон. Поскольку в клетках животных интроны обычно большие, а экзоны маленькие, существует высокая вероятность того, что случайный кусочек геномной ДНК будет иметь необходимую структуру: экзон, окруженный частичными последовательностями интронов. Вдобавок можно заметить, что улавливание экзонов моделирует события, происходившие в процессе эволюции генов (см. разд. 3.8 «Как эволюционировали прерывистые гены?»).

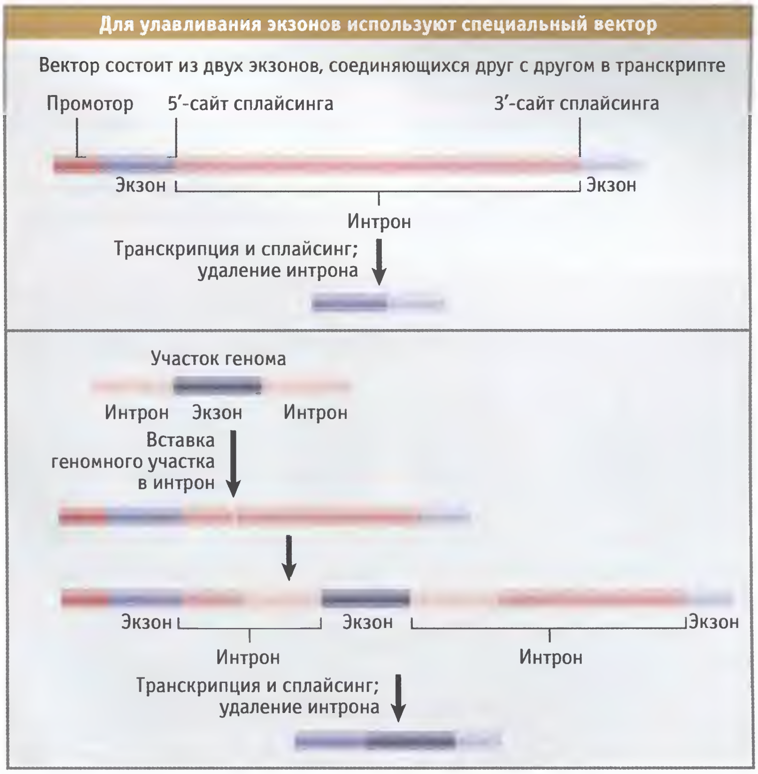


РИС. 4.11. Для улавливания экзонов используют специальный сплайсинг-вектор. Если в фрагменте геномной ДНК есть экзон, то его последовательность окажется в цитоплазматической РНК. Если же фрагмент генома состоит только из интрона, сплайсинг даст короткую РНК «пустого» вектора

4.8 Консервативность организации генома позволяет выявлять гены

Ключевые понятия:

- Компьютерные алгоритмы выявления генов неидеальны, и получаемые таким образом данные необходимо корректировать вручную.
- Псевдогены необходимо отличать от активных генов.
- Уровень синтении между человеческим и мышиным геномами очень высок, наиболее активные гены лежат в синтеничных регионах.

Теперь, когда мы научились определять полные последовательности геномов, наиболее актуальной становится задача локализации генов на последовательности нуклеотидов. К сожалению, кодирующие участки составляют лишь малую часть генома. Можно пытаться выявлять экзоны как сравнительно протяженные открытые рамки считывания, фланкированные характерными для экзон-интронных стыков последовательностями. Но, возможно, существуют дополнительные критерии, позволяющие выявить на «безликой» последовательности активный ген?

На РИС. 4.12 показано, что активный ген должен удовлетворять следующим критериям: 1) состоять из серии экзонов, первый из которых должен быть расположен сразу за промотором; 2) внутренние экзоны должны быть фланкированы характерными соединительными участками для сплайсинга (сайтами сплайсинга); 3) за последним экзоном должен идти сигнал полиаденилирования; и 4) объединять все эти экзоны вместе должна одна рамка считывания, начинающаяся старт-кодоном и заканчивающаяся стоп-кодоном. Внутренние экзоны можно локализовать как открытые рамки считывания, фланкированные сайтами сплайсинга. В простейшем случае первый и последний экзон включают, соответственно, начало и конец кодирующего региона (место его стыка с 5'- и 3'-нетранслируемыми регионами). В более сложных случаях первый и последний экзоны могут сами по себе быть нетранслируемыми регионами, и тогда выявить их значительно труднее.

Алгоритмы, используемые для локализации экзонов, недостаточно эффективны в случаях, когда гены очень велики, и экзоны расположены далеко друг от друга. Например, в результате начального анализа генома человека, на карту были нанесены 170 000 экзонов, входящих в состав 32 000 генов. Это, скорее всего, не соответствует истине, поскольку в таком случае на ген в среднем приходится 5,3 экзона, тогда как для полностью охарактеризованных на сегодняшний день генов это значение в среднем составляет 10,2 экзона на ген. Либо мы пропускаем большое количество экзонов, либо оставшиеся экзоны должны быть объединены в меньшее число генов.

Даже когда структура гена установлена верно, есть проблема с отделением активных генов от псевдогенов. Многие псевдогены можно распознать по явным дефектам (в виде множественных мутаций, приводящих к формированию неактивной кодирующей последовательности). Псевдогены, возникшие сравнительно недавно, еще не успели накопить достаточного количества мутаций, и их может быть сложно распознать. В качестве яркого примера можно привести мышиный ген *Gapdh* (кодирующий глицеральдегидфосфатдегидрогеназу). У мышей встречается лишь одна активная копия этого гена и около 400 псевдогенов. Около ста из них при первичном анализе генома выглядели активными, и для исключения их из списка активных генов потребовалась отдельная проверка каждого.

Дополнительным подтверждением того, что ген активен, может служить сравнение участков геномов у разных видов. Геномы человека и мыши претерпели значительную реорганизацию. Это видно хотя бы по тому, что гаплоидный геном человека состоит из 23 хромосом, а мыши — из 20. Однако на уровне отдельных реги-

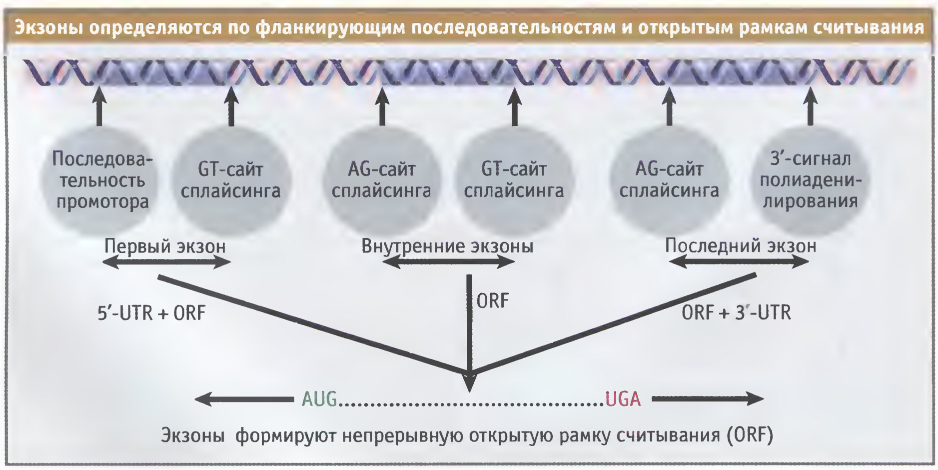


РИС. 4.12. Экзоны генов, кодирующих белки, определяются как кодирующие участки, фланкированные определенными сигнальными последовательностями (и нетранслируемыми участками, UTR, по обоим концам). Последовательности экзонов формируют открытые рамки считывания с соответствующими старт- и стоп-кодонами

онов порядок расположения генов в целом похож: при сравнении гомологичных участков геномов человека и мыши существует тенденция и к гомологии генов. Такую взаимосвязь называют **синтенией** (synteny).

На **РИС. 4.13** показана связь между первой хромосомой мыши и набором хромосом человека. В мышиной хромосоме мы можем обнаружить 21 сегмент, имеющий синтенных партнеров в хромосомах человека. Степень перетасовки, произошедшей в геномах, выражается в том, что эти сегменты разбросаны по шести различным хромосомам человека. Связи похожего типа можно найти во всех мышиных хромосомах, кроме X-хромосомы. Это объясняется тем, что X-хромосома — особый случай, ее структура во многом зависит от компенсаторных мер, необходимых для сглаживания различий между самцами (имеющими одну копию X-хромосомы) и самками (имеющими две копии) (см. разд. 31.5 «X-хромосомы претерпевают глобальные изменения»). Это повышает давление отбора, направленное против переноса генов в X-хромосому или из нее.

Сравнение геномов человека и мыши показывает, что >90% последовательностей каждого из геномов лежат в синтенных блоках, варьирующих по размеру (от 300 тпн до 65 Мпн). Всего обнаруживается 342 синтенных сегмента, средней длинной около 7 Мпн (0,3% от всего генома). 99% процентов мышиных генов имеют гомологов в геноме человека; 96% этих гомологов находятся в синтенных регионах.

Сравнение геномов раскрывает интересную особенность эволюции видов. Число семейств генов у человека и мыши одинаково, а основное различие между видами выражается в различной представленности определенных семейств в одном и другом геномах. Это особенно заметно в случаях с генами, влияющими на фенотипические признаки, специфичные для данного вида. Из 25 широко представленных семейств генов у мышей 14 содержат гены, ответственные за репродукцию грызунов, и 5 — гены, специфичные для их иммунной системы.

Подтверждением важности синтенных блоков является попарное сравнение содержащихся в них генов. При поиске возможных псевдогенов (основанном на сравнении последовательностей) ген, расположенный вне синтенного блока (то есть контекст которого отличается у двух видов), с примерно в два раза большей вероятностью является псевдогеном. Также и транслокация, расположенная вне исходного локуса, скорее всего, связана с возникновением псевдогена. Недостаток гомологичного гена в синтенной позиции подкрепляет уверенность в том, что наблюдаемый ген в действительности является псевдогеном. В целом >10% генов, выявленных в ходе первичного анализа генома, на самом деле являются псевдогенами.

Как правило, сравнение геномов делает процесс распознавания генов более эффективным. Когда последовательности особенно консервативны (например, как между мышью и человеком), вероятность того, что они являются активными гомологами, возрастает.



РИС. 4.13. Первая хромосома мыши состоит из 21 сегмента длиной от 1 до 25 Мпн, синтенных участкам шести хромосом человека

Определение генов, кодирующих структурные РНК, — более сложная задача, поскольку мы не можем воспользоваться критерием открытой рамки считывания. Однако и в этом случае сравнение геномов повышает точность анализа. Например, сравнение геномов человека и мыши обнаруживает около 500 генов, кодирующих тРНК, однако сравнение их позиционных особенностей позволяет сказать, что не более 350 из них в каждом гене являются активными.

4.9 В органеллах есть собственная ДНК

Ключевые понятия:

- Митохондрии и хлоропласты обладают собственными геномами, демонстрирующими неменделевское наследование. Обычно они наследуются по материнской линии.
- Геномы органелл растений могут претерпевать соматическую сегрегацию.
- Анализ митохондриальной ДНК показывает, что люди произошли от одной женщины, жившей в Африке 200 000 лет назад.

Первым доводом, говорящим в пользу наличия генов за пределами ядра, послужили данные о неменделевском наследовании признаков у растений (полученные в начале двадцатого века сразу после вторичного открытия законов Менделя). Неменделевское наследование иногда связывают с феноменом соматической сегрегации. Причина и у того, и у другого одна:

- Неменделевское наследование определяется отсутствием менделевского расщепления родительских признаков у потомка. Оно отражает недостаточную связь между характером расщепления и поведением хромосом в мейозе.
- Соматическая сегрегация описывает феномен, при котором родительские признаки расходятся по разным соматическим клеткам, формируя, таким образом, разнородность организма. Соматическая сегрегация — важная особенность развития растений, поскольку она отражает недостаточную сцепленность разделяющегося признака и поведения хромосом при митозе.

Таким образом, неменделевское наследование и соматическая сегрегация могут быть использованы для обнаружения генов, находящихся вне ядра и не подчиняющихся законам распределения гомологичных хромосом по гаметам и дочерним клеткам при помощи мейотического или митотического веретена деления. На **РИС. 4.14** показан пример, когда подобная ситуация возникает, если митохондрии, наследуемые от самца и самки, содержат различные аллели и дочерняя клетка случайно получает один из вариантов, формируя при этом несбалансированный набор митохондрий, соответствующих митохондриям одного из родителей (см. разд. 17.12 «Как происходит репликацияя и сегрегация ДНК в митохондриях эукариот?»).

Крайняя форма неменделевского наследования — *однородительское наследование*, когда наследуется генотип только одного из родителей, а генотип второго родителя постоянно теряется. В менее «категоричном» случае потомство с генотипом одного из родителей преобладает над потомством, наследующим другой генотип. Обычно предпочтительнее наследуется материнский генотип (либо наследуется только он). Это явление иногда описывают термином **материнское наследование**. Здесь важно то, что в потомстве преобладает генотип родителя одного пола, что выражается в аномальном разделении признаков при скрещивании мутанта и дикого типа. Эта ситуация контрастирует с законами менделевской генетики, когда при реципрокных скрещиваниях признаки обоих родителей наследуются в равной степени.

Отклонения в представленности родительских генотипов формируются во время или вскоре после формирования зиготы. Существует множество возможных причин таких отклонений. Вклад материнской и отцовской информации в органеллы зиготы может быть неравнозначным (а в некоторых случаях этот вклад осуществляет только один из родителей). Иногда вклады равны, но информация, предоставляемая одним из родителей, по каким-то причинам элиминируется. Встречаются также и комбинации перечисленных эффектов. Какой бы ни была причина «ненормального» наследования признаков, неравная представленность родительской информации контрастирует с равномерным распределением этой информации в ядре.

Причиной неменделевского наследования является наличие геномов в митохондриях и хлоропластах, наследуемых независимо от ядерных генов. По сути, геномы органелл представляют собой ДНК, физически изолированную в определенной части клетки и сформировавшую собственные механизмы экспрессии и регуляции. Геномы органелл могут кодировать некоторые или даже все собственные РНК, но в то же время кодируют только некоторые из белков, необходимых для поддержания жизнедеятельности органеллы. Остальные белки закодированы в ядерной ДНК и, экспрессируясь через цитоплазматический аппарат синтеза белков, импортируются в органеллы.

Гены, находящиеся вне ядра, называют **внеядерными**; они транскрибируются и транслируются в *той же* органелле (митохондрии или хлоропласте), в которой они расположены. Напротив, *ядерные* гены экспрессируются посредством *цитоплазматического* синтеза белков (иногда для описания поведения генов, расположенных в органеллах, используют термин **цитоплазматическая наследственность**. Мы не будем его использовать, поскольку важно разделять события, происходящие в цитоплазме и собственно в органеллах).

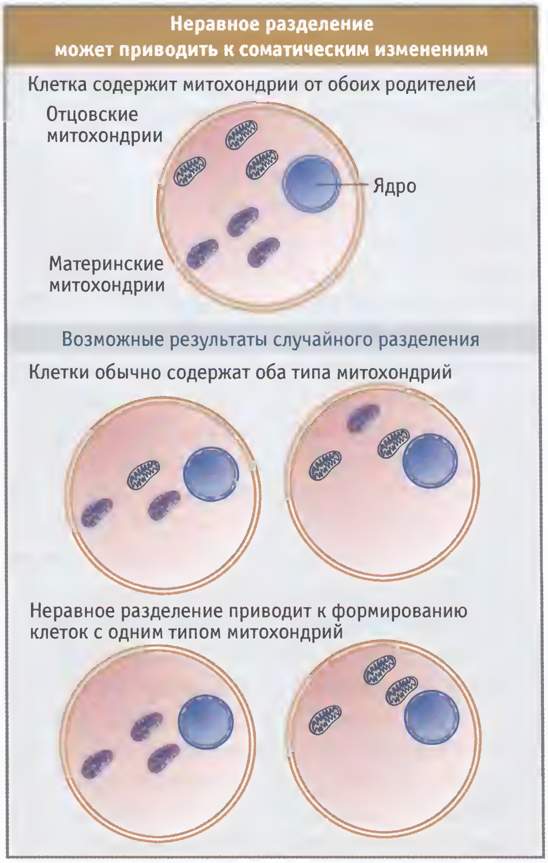


РИС. 4.14. Если аллели митохондриальной ДНК отца и матери отличаются, дочерняя клетка несет два набора митохондриальных ДНК. В результате митоза обычно формируется дочерняя клетка с обоими наборами. Если в результате неравного разделения формируется дочерняя клетка с одним набором, могут возникать соматические изменения

Высшие животные демонстрируют материнское наследование, объясняемое тем, что митохондрии передаются исключительно через яйцеклетки (и никогда не передаются со сперматозоидом). На **РИС. 4.15** показано, что вклад сперматозоида состоит только в передаче копии ядерной ДНК. Таким образом, митохондриальные гены передаются исключительно по материнской линии, а по отцовской в каждом следующем поколении они теряются.

Условия в органелле отличаются от ядерных, поэтому ДНК в органеллах эволюционирует по собственному пути. Если наследование однородительское, то рекомбинации между родительскими геномами быть не может. На самом деле рекомбинация не происходит и в тех

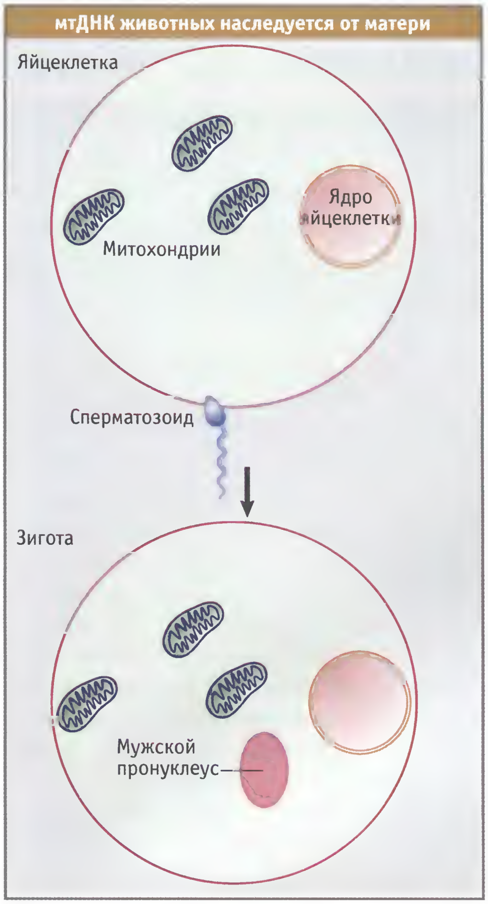


РИС. 4.15. ДНК, содержащаяся в сперматозоиде, проникает в яйцеклетку и формирует пронуклеус, однако все митохондрии предоставляются яйцеклеткой

случаях, когда геномы органелл наследуются от обоих родителей.

ДНК органелл имеет отличную от ядра систему репликации; в результате частота возникновения ошибок, совершаемых во время репликации, может также отличаться от ядерной. У млекопитающих митохондриальная ДНК накапливает мутации быстрее, чем ядерная, а у растений — медленнее (в хлоропластах скорость накопления мутаций демонстрирует среднее значение).

Одним из последствий материнского наследования является то, что последовательность митохондриальной ДНК более чувствительна к уменьшению популяции особей, способных к размножению, чем последовательность ядерной ДНК. Сравнение последовательностей митохондриальной ДНК на уровне человеческих популяций позволяет реконструировать распространение *Homo sapiens*. Расхождение между митохондриальными ДНК человека достигает 0,57%. Оказалось, что можно построить филогенетическое дерево, в котором варианты митохондрий расходятся от общего (африканского) предка. Скорость, с которой митохондриальная ДНК накапливает мутации, составляет 2–4% за миллион лет, что примерно в 10 раз быстрее, чем та же скорость для генов глобинов. Такая скорость мутирования позволяет датировать начало эволюционного периода в диапазоне от 140 000 до 280 000 лет назад. Это означает, что раса людей произошла от женщины, жившей в Африке примерно 200 000 лет назад.

4.10 Геномы органелл — кольцевые ДНК, в которых закодирована информация о некоторых белках органелл

Ключевые понятия:

- Геномы органелл обычно (но не всегда) представлены кольцевыми молекулами ДНК.
- Геномы органелл кодируют только некоторые из белков органелл.

Большинство геномов органелл представляют собой единственную кольцевую молекулу ДНК с уникальной последовательностью (обозначают как **мтДНК** в митохондриях и **хлДНК** в хлоропластах). Существует несколько исключений, для которых митохондриальная ДНК представлена не кольцевой, а линейной молекулой (такие случаи, как правило, характерны для низших эукариот).

Обычно в каждой органелле существует несколько копий генома, а поскольку в клетке присутствует много органелл данного типа, на одну клетку, в конечном счете, приходится большое количество одинаковых геномов. Таким образом, несмотря на то что сам по себе геном органеллы уникален, по частоте встречаемости в клетке он сопоставим с ядерной высокоповторяющейся последовательностью.

Геномы хлоропластов имеют сравнительно большие размеры — обычно около 140 тпн у высших растений и до 200 тпн у одноклеточных. Они сопоставимы по размеру с геномом крупного бактериофага (например, фага T4, геном которого достигает 165 тпн). На одну клетку обычно приходится множество копий хлоропластов (несколько десятков), и в каждой органелле содержится множество копий генома (обычно от 20 до 40 для высших растений).

Размеры полных митохондриальных геномов могут различаться более чем на порядок. Митохондриальные геномы животных малы — у млекопитающих они достигают примерно 16,5 тпн. В клетке обычно находится несколько сотен митохондрий, и каждая митохондрия содержит множество копий ДНК. Вместе с тем суммарное количество митохондриальной ДНК по сравнению с ядерной ДНК довольно мало — менее 1%.

Митохондриальный геном дрожжей значительно больше, чем у животных. Его размер у различных штаммов *S. cerevisiae* варьирует в широких пределах, составляя в среднем около 80 тпн. В дрожжевой клетке содержится около 22 митохондрий, каждая из которых содержит примерно 4 копии генома на органеллу. В рас-

тущих клетках доля митохондриальной ДНК может достигать 18% от суммарной ДНК клетки.

У растений размеры митохондриальной ДНК варьируют в широких пределах; ее минимальный размер составляет примерно 100 тпн. Довольно сложно выделить ДНК такого размера в неповрежденном виде, однако рестрикционное картирование, выполненное для нескольких растений, позволяет утверждать, что их митохондриальный геном представлен единственной кольцевой молекулой ДНК. В пределах одного генома встречаются короткие гомологичные последовательности. Рекомбинация между этими элементами приводит к выщеплению маленьких субгеномных кольцевых молекул, сосуществующих с полноразмерным «оригинальным» геномом. Это яркий пример сложности организации митохондриальной ДНК растений.

Располагая последовательностями митохондриальных генов многих организмов, мы можем представить основную схему функционирования митохондриальной ДНК. На **РИС. 4.16** приведен размер и генный состав нескольких митохондриальных геномов. Общее число генов, кодирующих белки, довольно мало и не коррелирует с размером генома. Геномы митохондрий млекопитающих (размером всего 16 тпн) кодируют 13 белков, тогда как в митохондриях дрожжей (содержащих геномы размером от 60 до 80 тпн) закодировано лишь восемь белков. У растений значительно большие по размеру митохондриальные геномы кодируют и большее число белков. Интроны есть в большинстве митохондриальных геномов (за исключением маленьких митогеномов млекопитающих).

Две ключевые рРНК всегда кодируются митохондриальными геномами. Количество тРНК, кодируемых митохондриальными геномами, варьирует от нуля до полного набора (в митохондриях 25–26). Это объясняет вариацию показателей на рис. 4.16.

Основной кодирующий потенциал митохондриальных геномов сосредоточен в регионах, содержащих гены ферментов дыхательных комплексов I–IV. У протист и растений в митохондриях закодировано множество рибосомальных белков, однако их мало или практически нет в митохондриальных геномах грибов и животных. У многих протист в митохондриях есть гены, кодирующие белки, вовлеченные в процессы импорта веществ в митохондрии.

Митохондриальная ДНК кодирует РНК и белки

Группы организмов	Размер (тпн)	Гены, кодирующие белки	Гены, кодирующие РНК
Грибы	19–100	8–14	10–28
Простейшие	6–100	3–62	2–29
Растения	186–366	27–34	21–30
Животные	16–17	13	4–24

РИС. 4.16. Митохондриальные геномы содержат гены, кодирующие белки (в основном комплексы I–IV), рРНК и тРНК

4.11 Митохондриальная ДНК может иметь различную структуру

Ключевые понятия:

- Митохондриальная ДНК животной клетки черезвычайно компактна и, как правило, кодирует всего 13 белков, 2 рРНК и 22 тРНК.
- Митохондриальные гены дрожжей содержат большие интроны, поэтому их митохондриальный геном в 5 раз длиннее, чем мтДНК животных.

Митохондриальная ДНК животных организована чрезвычайно компактно. При подробном рассмотрении организации митохондриальных генов различных групп животных обнаруживаются значительные различия, при этом основной принцип сохраняется: маленький геном кодирует ограниченное число функций. Митохондриальные геномы млекопитающих не отклоняются от этого правила. В их митохондриальных генах нет интронов, некоторые гены перекрываются, и почти каждая пара оснований принадлежит какому-либо гену. За исключением D-петли (региона, отвечающего за инициацию репликации ДНК), к межцистронным участкам можно отнести не более 87 из 16 569 пн митохондриального генома человека.

Полные нуклеотидные последовательности митохондриальных геномов животных демонстрируют значительное сходство. Карта митохондриального генома человека представлена на **РИС. 4.17**. На ней имеются 13 регионов, кодирующих белки. Все белки являются компонентами аппарата, связанного с клеточным дыханием. К ним относятся цитохром *b*, три субъединицы цитохромоксидазы, одна из субъединиц АТФазы и семь субъединиц (или связанных белков) NADH-дегидрогеназы.

Пятикратное различие в размерах митохондриальных геномов *S. cerivisiae* (84 тпн) и млекопитающих (16 тпн) предупреждает нас о том, что вопреки сходству выполняемых митохондриями функций в их генетической организации должны существовать значительные различия. Если судить по числу выполняемых митохондриями ферментативных функций, то количество эндогенно синтезируемых продуктов должно быть сходным. Какова же роль дополнительного генетического материала митохондрий дрожжей? Кодирует ли он другие белки, связан ли он с регуляторными процессами, или же он вовсе не экспрессируется?

Карта, представленная на **РИС. 4.18**, демонстрирует основные гены митохондрий дрожжей. Самой заметной отличительной чертой этой карты является удаленность кодирующих локусов.

Два наиболее заметных локуса содержат прерывистые гены *box* (кодирует цитохром *b*) и *oxi3* (кодирует первую субъединицу цитохром-оксидазы). Вместе эти два гена сравнимы по величине со всем митохондриальным геномом млекопитающих! Многие из длинных интронов этих генов содержат самостоятельные открытые рамки считывания (см. разд. 27.5 «Интроны I группы кодируют эндонуклеазы, поддерживающие их мобильность»). Это добавляет несколько белков (синтезируе-

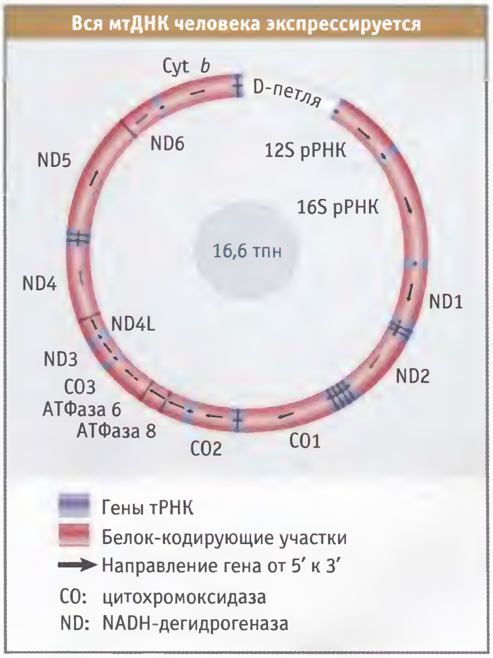


РИС. 4.17. Митохондриальная ДНК человека содержит 22 гена, кодирующих тРНК, две рРНК и 13 регионов, кодирующих белки. 14 из 15 регионов, кодирующих белок или рРНК, транскрибируются в одном направлении. 14 генов тРНК экспрессируются по часовой стрелке и 8 — против

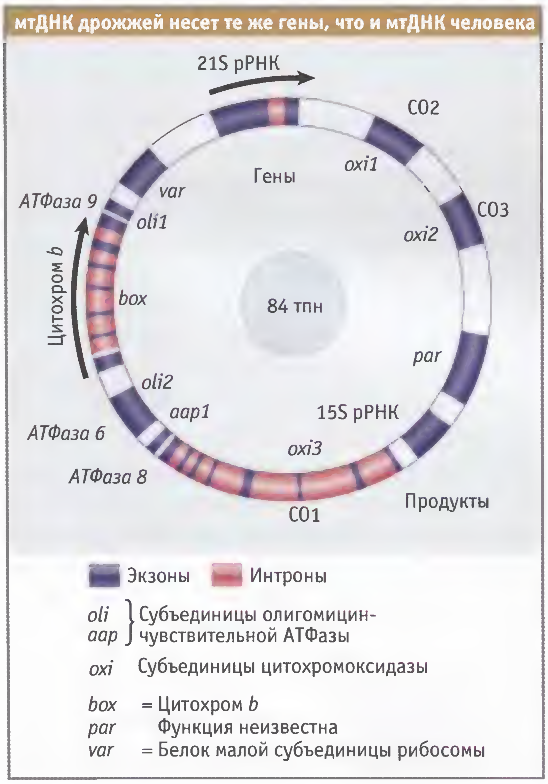


РИС. 4.18. Митохондриальный геном *S. cerivisiae* содержит как прерывистые, так и непрерывные гены, кодирующие белки, рРНК и тРНК (места расположения не обозначены). Стрелки указывают направление транскрипции

мых, впрочем, в незначительных количествах) к общему числу белков, кодируемых митохондриальным геномом дрожжей.

Остальные гены безынтронны. Они кодируют две оставшиеся митохондриальные субъединицы цитохромоксидазы, субъединицу (или субъединицы) АТФазы и (в случае с *var1*) митохондриальный рибосомальный белок. Общее число митохондриальных генов не превышает 25.

4.12 Геномы хлоропластов кодируют сравнительно много белков и РНК

Ключевые понятия:

- Геномы хлоропластов варьируют по размеру, однако они достаточно велики для того, чтобы кодировать от 50 до 100 белков, а также рРНК и тРНК.

Какие же гены находятся в геноме хлоропластов? ДНК хлоропластов варьирует по длине от 120 до 190 тпн. Хлоропластные геномы с известной полной нуклеотидной последовательностью (известно более 10 геномов) содержат от 87 до 183 генов. На РИС. 4.19 перечислены функции, кодируемые геномами хлоропластов наземных растений. Для геномов хлоропластов водорослей отмечена более значительная вариация показателей.

В целом ситуация напоминает таковую для митохондрий, за исключением того, что общее число генов больше. Геномы хлоропластов кодируют все типы рРНК и тРНК, необходимые для собственной (хлоропластной) системы синтеза белка. Помимо обычных, рибосома содержит две дополнительные маленькие рРНК. ДНК хлоропластов может содержать все необходимые гены для синтеза полного набора тРНК. Помимо этого, геномы

В ДНК хлоропластов более 100 генов

Гены	Варианты
Кодирующие РНК	
16S рРНК	1
23S рРНК	1
4,5S рРНК	1
5S рРНК	1
тРНК	30–32
Кодирующие гены	
Белки рибосом	20–21
РНК-полимераза	3
Другие	2
Функции, свойственные хлоропластам	
Рубиско (рибулозобисфосфат-карбоксилаза-оксигеназа) и тилакоиды	31–32
NADH-дегидрогеназа	11
Всего	105–113

РИС. 4.19. Геном хлоропластов наземных растений кодирует 4 рРНК, 30 тРНК и около 60 белков

хлоропластов кодируют около 50 белков, включая РНК-полимеразу и рибосомальные белки. Правило, что гены органелл транскрибируются и транслируются самой органеллой, справедливо и для хлоропластов.

Около половины генов хлоропластов кодируют белки, вовлеченные в процесс синтеза белков. В качестве подтверждения эндосимбиотического происхождения хлоропластов можно отметить гомологичность их генов к генам бактерий. В частности, организация генов рРНК очень схожа с таковой у цианобактерий, что позволяет точно определить последнего общего предка между хлоропластами и бактериями.

Интроны хлоропластов делят на два основных класса. Те из них, которые находятся в генах тРНК, обычно расположены в антикодоновой петле (наподобие интронов в генах ядерной тРНК дрожжей (см. разд. 26.14 «Сплайсинг тРНК дрожжей идет через разрезание и сшивание предшественника)). Те же, что расположены в генах, кодирующих белки, напоминают интроны митохондриальных генов (см. гл. 27 «Каталитическая РНК»). Таким образом, в эволюции эндосимбиотическое возникновение хлоропластов относится ко времени, предшествовавшему появлению прокариот с безынтронными генами.

Функция хлоропластов — осуществление фотосинтеза. Многие из их генов кодируют белковые комплексы мембран тилакоидов. Ферментативный состав этих комплексов отличается от ключевых ферментов митохондрий. Сходство хлоропластов и митохондрий можно заметить и в том, что некоторые из субъединиц сложных белковых комплексов хлоропластов, так же, как комплексы митохондрий, частично кодируются геномом органеллы, а частично — ядерным геномом. Другие же белковые комплексы целиком кодируются геномом либо хлоропласта, либо ядра.

4.13 Митохондрии возникли в результате эндосимбиоза

Как получилось, что органелла несет информацию, необходимую для осуществления лишь некоторых из ее функций, в то время как остальные функции закодированы в ядре? На **РИС. 4.20** показана эндосимбиотическая модель эволюции митохондрий (теория симбиогенеза), по которой митохондрии и хлоропласты возникли в эволюции из примитивных клеток, захваченных бактериями. С этой точки зрения протоорганелла должна была содержать все гены, необходимые для осуществления ее функций.

Сравнение последовательностей геномов митохондрий и хлоропластов заставляет думать, что они эволюционировали начиная от неких предков, имевших общие черты с эубактериями. Митохондрии имеют общее происхождение с α-пурпурными бактериями, а хроропласты делят происхождение с цианобактериями. Из бактерий

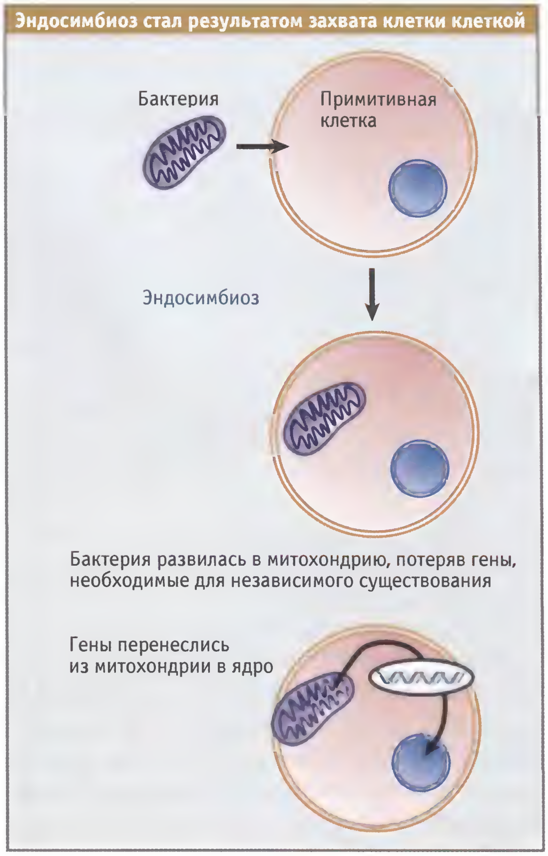


РИС. 4.20. Митохондрии возникли в результате захвата бактерии эукариотической клеткой (эндосимбиоз)

ближайшим известным родственником митохондрий является *Rickettsia* (возбудитель тифа) — облигатный внутриклеточный паразит, вероятно, происходящий от свободноживущих бактерий. Этот пример подтверждает теорию эндосимбиотического возникновения митохондрий от предка, также имевшего общие черты с *Rickettsia*.

Для того чтобы бактерия «встроилась» в клетку-реципиент и эволюционировала в митохондрию (или хлоропласт), ей необходимо было претерпеть два значительных изменения. В органеллах гораздо меньше генов, чем в независимой бактерии, таким образом, они должны были утерять некоторые гены, ответственные за выполнение функций, необходимых для независимого существования (например, гены компонентов метаболических путей). На деле большинство генов, кодирующих функции органелл, расположены в ядре. Следовательно, их нужно было как-то перенести туда из органеллы.

Перенос ДНК между органеллой и ядром возник в процессе эволюции и продолжается до сих пор. Скорость переноса может быть измерена напрямую путем вставки в органеллу гена, способного функционировать только в ядре. Например, такой ген может содержать специфичный для ядра интрон, или он может кодировать белок, функционирующий только в цитоплазме. Говоря о скорости переноса генетического материала как о факторе эволюции, следует отметить, что она в точнос-

ти соответствует скорости изменения одного гена. ДНК митохондрий переносится в ядро со скоростью около 2×10^{-5} за поколение. Эксперименты по определению скорости переноса в обратном направлении, из ядра в митохондрию демонстрируют значительно меньшую скорость — менее 10^{-10} за поколение. Когда способный работать только в ядре ген устойчивости к антибиотику встраивают в геном хлоропласта, его перенос и успешную экспрессию можно определить, исследуя рассаду на устойчивость к данному антибиотику. Подобные эксперименты показывают, что перенос осуществляется у 1 из 16 000 потомков, что соответствует скорости 6×10^{-5}.

Перенос гена из органеллы в ядро, конечно, требует физического передвижения ДНК, но для успешной экспрессии необходимы еще и изменения в кодирующей последовательности. Белки органелл, кодируемые ядерными генами, содержат специальные последовательности, позволяющие импортировать их в органеллу после того, как они были синтезированы в цитоплазме (см. разд. 10.16 «Посттрансляционная интеграция в мембрану зависит от лидерных последовательностей»). Белкам, синтезируемым в органелле, такие последовательности не нужны. Возможно, наиболее эффективный перенос генов происходил в то время, когда части клетки не были так обособлены. Поэтому как перемещение ДНК в ядро, так и встраивание белков в органеллы могло происходить гораздо легче.

Филогенетические деревья демонстрируют, что перенос генов происходил независимо в различных линиях живых организмов. Судя по всему, перенос митохондриальных генов в ядро происходил на ранней стадии эволюции животной клетки, а в клетках растений этот процесс, возможно, продолжается до сих пор. Число транслокаций генов может быть очень большим. Так, у *Arabidopsis* более 800 ядерных генов гомологичны (по последовательности) генам хлоропластов других растений. Следовательно, это могут быть гены, берущие свое начало из хлоропластов.

4.14 Резюме

Последовательности ДНК, составляющие эукариотические геномы, можно разделить на три группы:

- уникальные (неповторяющиеся) последовательности;
- умеренно повторяющиеся последовательности, разбросанные по геному и представленные в небольшом количестве в виде гомологичных, но не идентичных копий;
- короткие и представленные в виде тандемных структур, высокоповторяющиеся последовательности.

Соотношение этих типов последовательностей является характерной чертой каждого генома, хотя для более крупных геномов существует тенденция к возрастанию доли высокоповторяющихся ДНК. Почти 50% генома человека состоит из повторяющихся последовательностей, подавляющее большинство которых соответствует последовательностям транспозонов. Большая же часть генов представляет собой уникальные участки ДНК. Суммарная длина уникальных последовательностей коррелирует со сложностью организма лучше, чем общий размер генома; максимальная суммарная длина уникальных регионов в геноме составляет около 2×10^9 пн.

Неменделевское наследование объясняется наличием ДНК в расположенных в цитоплазме органеллах. Митохондрии и хлоропласты — это окруженные мембранами структуры, в пределах которых синтезируются лишь некоторые необходимые для органелл белки, в то время как остальные импортируются из цитоплазмы. Геномы органелл обычно представлены кольцевой молекулой ДНК, кодирующей, как правило, все рРНК, тРНК, и некоторые необходимые для существования органеллы белки.

Размеры геномов митохондрий варьируют в широком диапазоне: от 16 тпн у млекопитающих до 570 тпн у высших растений. Более крупные геномы органелл могут кодировать больше функций. Геномы хлоропластов варьируют в пределах 120–200 тпн. Те из них, для которых определена полная нуклеотидная последовательность, демонстрируют схожую структуру и набор кодируемых функций. Субъединицы основных белков митохондрий и хлоропластов могут синтезироваться как в пределах органеллы, так и импортироваться из цитоплазмы.

У дрожжей довольно часто возникают перестановки ДНК, причем рекомбинация обнаружена как для митохондриальных геномов, так и для геномов хлоропластов. Между хлоропластами и митохондриями, с одной стороны, и ядром, с другой, может происходить перенос генетического материала.

Литература

4.3 Геномы отличаются огромным разнообразием

Исследование

Altshuler, D., Brooks, L. D., Chakravarti, A., Collins, F. S., Daly, M. J., and Donnelly, P. (2005). A haplotype map of the human genome. *Nature* 437, 1299–1320.

Altshuler, D., Pollara, V. J., Cowles, C. R., Van Etten, W. J., Baldwin, J., Linton, L., and Lander, E. S. (2000). An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature* 407, 513–516.

Mullikin, J. C., Hunt, S. E., Cole, C. G., Mortimore, B. J., Rice, C. M., Burton, J., Matthews, L. H., Pavitt, R., Plumb, R. W., Sims, S. K., Ainscough, R. M., Attwood, J., Bailey, J. M., Barlow, K., Bruskiewich, R. M., Butcher, P. N., Carter, N. P., Chen, Y., and Clee, C. M. (2000). An SNP map of human chromosome 22. *Nature* 407, 516–520.

4.4 ПДРФ и ОНП можно использовать для генетического картирования

Обзор

Gusella, J. F. (1986). DNA polymorphism and human disease. *Annu. Rev. Biochem.* 55, 831–854.

White, R., Leppert, M., Bishop, D. T., et al. (1985). Construction of linkage maps with DNA markers for human chromosomes. *Nature* 313, 101–105.

Исследование

Altshuler, D., Brooks, L. D., Chakravarti, A., Collins, F. S., Daly, M. J., and Donnelly, P. (2005). A haplotype map of the human genome. *Nature* 437, 1299–1320.

Dib, C., Faure, S., Fizames, C., et al. (1996). A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* 380, 152–154.

Dietrich, W. F., Miller, J., Steen, R., et al. (1996). A comprehensive genetic map of the mouse genome. *Nature* 380, 149–152.

Donis-Keller, J., Green, P., Helms, C., et al. (1987). A genetic linkage map of the human genome. *Cell* 51, 319–337.

Hinds, D. A., Stuve, L. L., Nilsen, G. B., Halperin, E., Eskin, E., Ballinger, D. G., Frazer, K. A., and Cox, D. R. (2005). Whole-genome patterns of common DNA variation in three human populations. *Science* 307, 1072–1079.

Sachidanandam, R., Weissman, D., Schmidt, S., et al. (2001). A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. The International SNP Map Working Group. *Nature* 409, 928–933.

4.5 Почему геномы такие большие?

Обзор

Gall, J. G. (1981). Chromosome structure and the C-value paradox. *J. Cell Biol.* 91, 3s–14s.

Gregory, T. R. (2001). Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. *Biol. Rev. Camb. Philos. Soc.* 76, 65–101.

4.6 Эукариотические геномы состоят из уникальных и повторяющихся последовательностей

Обзор

Britten, R. J. and Davidson, E. H. (1971). Repetitive and nonrepetitive DNA sequences and a speculation on the origins of evolutionary novelty. *Q. Rev. Biol.* 46, 111–133.

Davidson, E. H. and Britten, R. J. (1973). Organization, transcription, and regulation in the animal genome. *Q. Rev. Biol.* 48, 565–613.

4.7 Гены можно локализовать, используя консервативность их экзонов

Исследование

Buckler, A. J., Chang, D. D., Graw, S. L., Brook, J. D., Haber, D. A., Sharp, P. A., and Housman, D. E. (1991). Exon amplification: a strategy to isolate mammalian genes based on RNA splicing. *Proc. Natl. Acad. Sci. USA* 88, 4005–4009.

Kunkel, L. M., Monaco, A. P., Middlesworth, W., Ochs, H. D., and Latt, S. A. (1985). Specific cloning of DNA fragments absent from the DNA of a male patient with an X chromosome deletion. *Proc. Natl. Acad. Sci. USA* 82, 4778–4782.

Monaco, A. P., Bertelson, C. J., Middlesworth, W., Colletti, C. A., Aldridge, J., Fischbeck, K. H., Bartlett, R., Pericak-Vance, M. A., Roses, A. D., and Kunkel, L. M. (1985). Detection of deletions spanning the Duchenne muscular dystrophy locus using a tightly linked DNA segment. *Nature* 316, 842–845.

4.9 В органеллах есть собственная ДНК

Исследование

Cann, R. L., Stoneking, M., and Wilson, A. C. (1987). Mitochondrial DNA and human evolution. *Nature* 325, 31–36.

4.10 Геномы органелл — кольцевые ДНК, в которых закодирована информация о некоторых белках органелл

Обзор

Lang, B. F., Gray, M. W., and Burger, G. (1999). Mitochondrial genome evolution and the origin of eukaryotes. *Annu. Rev. Genet.* 33, 351–397.

4.11 Митохондриальная ДНК может иметь различную структуру

Обзор

Attardi, G. (1985). Animal mitochondrial DNA: an extreme example of economy. *Int. Rev. Cytol.* 93, 93–146.

Boore, J. L. (1999). Animal mitochondrial genomes. *Nucleic Acids Res.* 27, 1767–1780.

Clayton, D. A. (1984). Transcription of the mammalian mitochondrial genome. *Annu. Rev. Biochem.* 53, 573–594.

Gray, M. W. (1989). Origin and evolution of mitochondrial DNA. *Annu. Rev. Cell Biol.* 5, 25–50.

Исследование

Anderson, S., Bankier, A. T., Barrell, B. G., et al. (1981). Sequence and organization of the human mitochondrial genome. *Nature* 290, 457–465.

4.12 Геномы хлоропластов кодируют сравнительно много белков и РНК

Обзор

Palmer, J. D. (1985). Comparative organization of chloroplast genomes. *Annu. Rev. Genet.* 19, 325–354.

Shimada, H. and Sugiura, M. (1991). Fine structural features of the chloroplast genome: comparison of the sequenced chloroplast genomes. *Nucleic Acids Res.* 11, 983–995.

Sugiura, M., Hirose, T., and Sugita, M. (1998). Evolution and mechanism of translation in chloroplasts. *Annu. Rev. Genet.* 32, 437–459.

4.13 Митохондрии возникли в результате эндосимбиоза

Обзор

Lang, B. F., Gray, M. W., and Burger, G. (1999). Mitochondrial genome evolution and the origin of eukaryotes. *Annu. Rev. Genet.* 33, 351–397.

Исследование

Adams, K. L., Daley, D. O., Qiu, Y. L., Whelan, J., and Palmer, J. D. (2000). Repeated, recent and diverse transfers of a mitochondrial gene to the nucleus in flowering plants. *Nature* 408, 354–357.

Arabidopsis Initiative (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature* 408, 796–815.

Huang, C. Y., Ayliffe, M. A., and Timmis, J. N. (2003). Direct measurement of the transfer rate of chloroplast DNA into the nucleus. *Nature* 422, 72–76.

Thorsness, P. E. and Fox, T. D. (1990). Escape of DNA from mitochondria to the nucleus in S. cerevisiae. *Nature* 346, 376–379.

Последовательности генома и число генов

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

5.1 Введение

С тех пор, как в 1995 г. были впервые определены полные нуклеотидные последовательности бактериальных геномов, скорость секвенирования и длина определяемых за один подход последовательностей сильно увеличились. Первыми геномами, для которых была определена полная нуклеотидная последовательность, были сравнительно небольшие бактериальные геномы (не превышающие 2 Мпн). К 2002 г. была определена последовательность генома человека длиной 3000 Мпн. На сегодняшний день определены последовательности геномов широкого спектра организмов, включая бактерии, археи, дрожжи, низших эукариот, растения и животных, для которых известны полные последовательности геномов некоторых червей, мух, грызунов и млекопитающих.

Возможно, наиболее ценной для исследователя информацией, заключенной в последовательности генома, является количество генов. Из всех известных геномов наименьшим (470 генов) обладает облигатная внутриклеточная бактерия *Mycoplasma genitalium.* У свободноживущих бактерий число генов варьирует в диапазоне от 1700 до 7500. Примерно на том же уровне находятся геномы архей. Самый маленький (из известных) геном одноклеточных эукариот представлен 5300 генами. У червей и мух примерно по 18 500 и 13 500 генов соответственно, однако максимальных значений число генов достигает у млекопитающих (порядка 25000).

На **РИС. 5.1** приведены минимальные значения числа генов для шести групп организмов. Около 500 генов необходимо для существования самой простой (паразитической) клетки, примерно 1500 — для свободноживущего одноклеточного организма (бактерии), более 5000 — для эукариотической клетки, более 10 000 — для многоклеточного организма и более 13 000 — для организма с нервной системой. Конечно, многие виды содержат большее число генов, нежели приведенный минимум, необходимый для существования организма данного типа. Даже у близкородственных организмов число генов может сильно различаться.

В геномах бактерий и низших эукариот подавляющее большинство генов уникальны. Однако у высших эукариот многие гены образуют семейства, состоящие из близкородственных последовательностей. Конечно, некоторые эукариотические гены уникальны (формально их можно считать семейством с единственным представителем), однако многие принадлежат к семействам, состоящим из десяти и более членов. Число разных семейств значительно лучше коррелирует с общей сложностью организма, чем число генов в геноме.

Однако наиболее ценную информацию можно получить при сравнении собственно нуклеотидных последовательностей геномов. Теперь, когда известны последовательности геномов человека и шимпанзе, можно попытаться найти ответ на вопрос, что же делает человека уникальным.



РИС. 5.1 Минимальное число генов, требуемое для обеспечения всех жизненных функций, возрастает с увеличением сложности организма. Фото внутриклеточной паразитической бактерии любезно предоставлено Gregory P. Henderson and Grant J. Jensen, California Institute of Technology. Фото свободноживущей бактерии любезно предоставлено Karl O. Stetter, Universität Regensburg. Фото одноклеточных эукариот любезно предоставлено Eishi Noguchi, Drexel University College of Medicine. Фото многоклеточного эукариотического организма любезно предоставлено Carolyn B. Marks and David H. Hall, Albert Einstein College of Medicine, Bronx, NY. Фото высших растений любезно предоставлено Keith Weller/USDA. Фото млекопитающего © Photodisc.

5.2 Число генов у разных бактерий варьирует в десятки раз

На сегодняшний день проделана значительная работа по определению полных последовательностей геномов различных организмов. Список некоторых секвени-

рованных геномов приведен на **РИС. 5.2**. Он начинается с микоплазмы, у которой геном составляет лишь $0{,}6 \times 10^6$ пн и заканчивается геномом человека, состоящим из $3{,}3 \times 10^9$ пн. Между ними расположены такие важные экспериментальные животные, как, например, дрожжи, плодовая мушка и нематода.

По последовательностям геномов бактерий и архей видно, что практически вся их ДНК (обычно 85–90%) кодирует РНК и белки. На **РИС. 5.3** показано, что размеры их геномов варьируют в пределах одного порядка и пропорциональны числу генов в геноме. Длина среднего гена у них составляет около 1000 пн.

Все бактерии, размер генома которых не превышает 1,5 Мпн, являются облигатными внутриклеточными паразитами — они живут в пределах эукариотической клетки, которая обеспечивает их необходимыми промежуточными метаболитами. Их геном выполняет минимальное число функций, необходимых для построения клетки. По сравнению с бактериями, имеющими более крупные геномы, все классы генов у паразитических бактерий представлены в меньшем объеме, однако наиболее значимой редукции подвергаются локусы, кодирующие ферменты метаболических путей (поскольку продукты этих ферментов, в основном, предоставляются клеткой-хозяином), а также локусы, связанные с регуляцией экспрессии генов. Самый маленький геном у *Mycoplasma genitalium* (состоит приблизительно из 470 генов).

Археи обладают биологическими особенностями, позволяющими расположить их между прокариотами и эукариотами, но по размеру геномов и числу генов они попадают в одну группу с бактериями. Размеры их геномов варьируют в диапазоне от 1,5 до 3 Мпн и содержат 1500–2700 генов. *Methanococcus jannaschii* — это архебактерия, способная производить метан и живущая в условиях повышенного давления и температуры. По числу генов она похожа на *Haemophilus influenzae*, однако процент генов, гомологичных генам других организмов, довольно низок. Аппарат генной экспрессии *M. jannaschii* напоминает скорее эукариотический, чем прокариотический тип, в то время как аппарат клеточного деления больше похож на таковой у прокариот.

Археи и самые маленькие из свободно живущих бактерий содержат наименьшее число генов среди свободноживущих клеток. Самый маленький геном архей представлен примерно 1500 генами. Самый же маленький геном свободноживущих бактерий принадлежит термофильной *Aquifex aeolicus* (порядка 1,5 Мпн и 1512 генов). Геном «типичной» грамотрицательной бактерии *H. influenzae* содержит 1743 гена, средняя длина которых примерно 900 пн. Из приведенных примеров можно сделать вывод, что для существования свободноживущей клетки необходимо как минимум 1500 генов.

Размеры геномов бактерий могут отличаться на порядок. Минимальное и максимальное значения для них составляют: 0,6 Мпн и — чуть менее 8 Мпн. Протяженные геномы, как правило, содержат и больше генов. Самыми большими (из известных) прокариотическими геномами обладают *Sinorhizobium meliloti* и *Mesorhizobium loti* — азотфиксирующие бактерии, живущие на корнях растений. Размеры их геномов (более 7 Мпн и число генов (более 7500) сравнимы с таковыми у дрожжей.

Размер генома *E.coli* занимает среднее положение в ряду бактериальных геномов. Обычный лабораторный штамм кишечной палочки содержит 4288 генов, средняя

В секвенированных геномах насчитывается от 470 до 30 000 генов

Организм	Размер генома, Мпн	Число генов	Число локусов, по которым известны летальные мутации
Mycoplasma genitalium	0,58	470	~300
Rickettsia prowazekii	1,11	834	
Haemophilus influenzae	1,83	1 743	
Methanococcus jannaschi	1,66	1 738	
B. subtilis	4,2	4100	
E. coli	4,6	4 288	1800
S. cerevisiae	13,5	6 034	1090
S. pombe	12,5	4 929	
A. thaliana	119	25 498	
O. sativa (рис)	466	~40 000	
D. melanogaster	165	13 601	3100
C. elegans	97	18 424	
H. sapiens	3300	~25 000	

РИС. 5.2 Для ряда организмов размер генома и число генов определены с высокой точностью — на основании полных последовательностей их геномов. Число локусов с летальными мутациями оценивается по данным генетических исследований

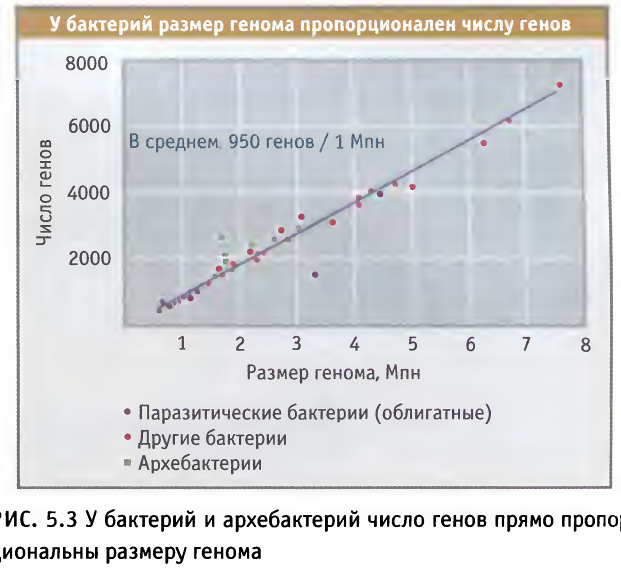


РИС. 5.3 У бактерий и архебактерий число генов прямо пропорциональны размеру генома

длина которых составляет около 950 пн, а промежутков между ними — 118 пн. Тем не менее штаммы могут сильно различаться. Максимальные и минимальные значения для *E. coli* варьируют в пределах от 4,6 Мпн и 4249 генов до 5,5 Мпн и 5361 ген.

До сих пор мы не знаем функций всех бактериальных генов. Для большей части геномов прокариот около 60% генов могут быть определены на основании гомологии с уже известными генами других видов. Эти гены делятся на несколько примерно равных классов: кодирующие белки метаболизма, структурные белки клетки (или компоненты транспортных систем), и регуляторные белки. Однако практически в каждом геноме есть около четверти генов, которым мы не можем приписать никакой функции на основании сравнения с базами данных. При этом большую часть этих генов можно найти в родственных организмах, что свидетельствует о консервативности выполняемой ими функции.

В связи с медицинской значимостью особое внимание уделяется определению последовательности геномов патогенных бактерий. Важным вкладом в развитие представлений о природе патогенности стало открытие того, что характерной чертой их геномов являются так называемые «островки патогенности». Это протяженные регионы (порядка 10–200 тпн), присутствующие в геномах патогенных видов бактерий и отсутствующие в геномах непатогенных штаммов тех же или родственных им видов. Нуклеотидный (G+C) состав этих регионов обычно отличается от остального генома, и, скорее всего, они могут мигрировать между бактериями по механизму горизонтального переноса. Так, например, бактерия, вызывающая сибирскую язву (*Bacillus anthracis*), содержит две больших плазмиды (экстрахромосомные кольцевые молекулы ДНК), каждая из которых несет островок патогенности, включающий гены, кодирующие бактериальный токсин.

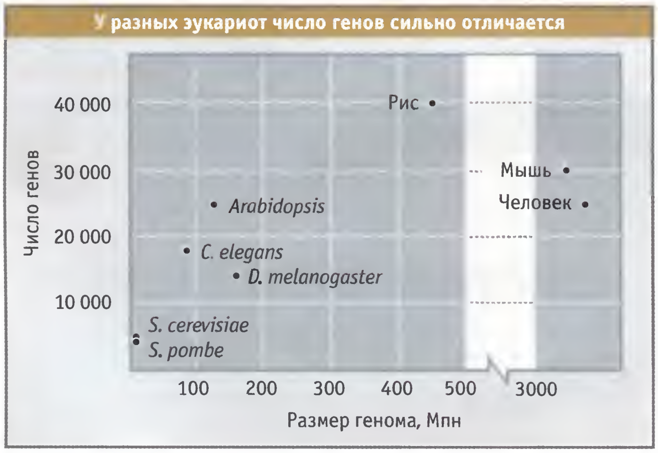


РИС. 5.4 Число генов в геномах эукариот принимает значения от 6000 до 40 000 вне корреляции с размером генома или сложностью организма

5.3 Для нескольких эукариотических организмов известно общее число генов

Ключевые понятия:

- У дрожжей — 6000 генов, у червей — 18 500, у мух — 13 600, у маленького растения *Arabidopsis* — 25 000, у мышей и людей, вероятно, — от 20 000 до 25 000.

Как уже было сказано, связь между размером генома и числом генов для геномов эукариот отсутствует. Геномы одноклеточных эукариот по размеру попадают в одну группу с самыми большими бактериальными геномами. Как можно видеть из РИС. 5.4, у высших эукариот больше генов, но их число не коррелирует с размерами генома.

Наибольшее количество данных по низшим эукариотам получено в результате изучения последовательностей геномов дрожжей *Saccharomyces cerevisiae* и *Schizosaccharomyces pombe*. На РИС. 5.5 представлены их отличительные черты. Геномы этих дрожжей (размером 12,5 и 13,5 Мпн) содержат около 6000 и около 5000 генов соответственно. Длина средней открытой рамки считывания составляет около 1,4 тпн. Получается, что около 70% генома этих дрожжей занято кодирующими регионами. Основное отличие между ними заключается в том, что

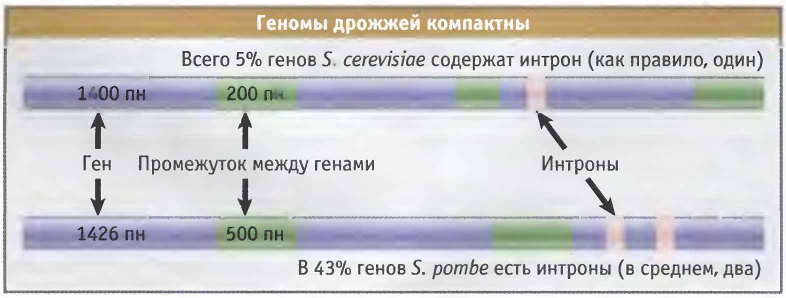


РИС. 5.5 Геном *S. cerevisiae,* размером 13,5 Мпн содержит 6000 генов, среди которых почти нет прерывистых. Геном *S. pombe,* размером 12,5 Мпн содержит 5000 генов: почти половина из них имеет интроны. *S. cerevisiae* и *S. pombe* обладают сходными размерами и генов, и промежутков между генами

лишь 5% генов *S. cerevisiae* (по сравнению с 43% *S. pombe*) содержат интроны. Плотность расположения генов высока, а их организация в целом сходна, хотя у *S. cerevisiae* расстояние между генами чуть меньше. Около половины обнаруженных в результате полногеномного секвенирования генов либо были известны ранее, либо оказались родственны известным генам. Оставшиеся гены — новые, и это дает некоторое представление о числе новых типов генов, которые могут быть открыты.

Выявление длинных открытых рамок считывания на основании анализа нуклеотидной последовательности — достаточно точный метод определения генов. Однако рамки считывания, кодирующие <100 аминокислот, не могут быть определены только на основании отсутствия стоп-кодонов (из-за частого возникновения ложноположительных результатов). Анализ экспрессии генов показывает, что у *S. cerevisiae* примерно 300–600 таких рамок являются истинными генами.

Достаточно эффективным способом подтверждения структуры гена является сравнение последовательностей у представителей близкородственных видов — если ген активный, он, скорее всего, будет консервативным. Сравнение последовательностей четырех близких видов дрожжей демонстрирует, что 503 гена, обнаруженных у *S. cerevisiae*, не имеют партнеров в геномах других видов и, таким образом, должны быть исключены из списка активных генов. Для *S. cerevisiae* это приводит к уменьшению общего числа генов до 5726.

Последовательность генома *Caenorhabditis elegans* представлена как богатыми на гены регионами, так и регионами, в которых гены встречаются редко. Полная последовательность содержит около 18 500 генов. Только 42% из них имеют предполагаемых гомологов за пределами класса нематод.

Несмотря на то что геном мух по размеру больше генома червей, в нем содержится меньшее число генов (около 13 600 у *D. melanogaster*). Вместе с тем число разных транкриптов у них достигает 14 100, что объясняется альтернативным сплайсингом. Неясно, почему муха — казалось бы, более сложно устроенный организм — содержит лишь 70% от общего числа генов у червей. Это еше раз подчеркивает отсутствие четкой взаимосвязи между числом генов и сложностью организации.

Геном растения *Arabidopsis thaliana* по размеру занимает промежуточное положение между червем и мухой, но при этом содержит большее число генов, чем любой из них (порядка 25000). Это также подтверждает отсутствие связи между числом генов и сложностью организма и подчеркивает особое свойство растений, у которых может быть больше генов (вследствие дупликаций у предков), чем в животных клетках. Большая часть генома *Arabidopsis* находится в дуплицированных сегментах, что свидетельствует о древнем удвоении генома (для возникновения тетраплоидности). Лишь 35% генов *Arabidopsis* представлены единичными копиями.

Геном риса (*Oryza sativa*) примерно в 4 раза больше, чем у *Arabidopsis,* тем не менее по числу генов он превышает его лишь на 50% (вероятно, около 40 000). Повторяющиеся последовательности занимают 42–45% генома. Более 80% генов, найденных у *Arabidopsis*, встречаются и в геноме риса. Из этих общих генов около 8000 встречаются у *Arabidopsis* и риса, но не встречаются ни в одном из известных нерастительных геномов. Вероятно, этот набор генов кодирует функции, свойственные только растениям (например, фотосинтез).

На примере генома мухи можно получить представление о том, сколько генов составляют отдельные функциональные классы белков. На **РИС. 5.6** существующие функциональные классы белков разбиты на несколько категорий. Среди известных генов мы находим 2500 генов, кодирующих ферменты, около 750, кодирующих транскрипционные факторы, примерно 700, кодирующих белки-транспортеры и белки ионных каналов и порядка 700, кодирующих белки, вовлеченные в передачу сигналов. Около половины генов кодируют продукты, выполняющие пока неизвестные функции. Примерно 20% белков расположены в мембранах.

Размеры белков увеличиваются в ряду от прокариот и архей к эукариотам. Так, у археи *M. jannaschii* и бактерии *E. coli* средняя величина белков составляет 287 и 317 аминокислот соответственно, тогда как у *S. cerivisiae* и *C. elegans* средняя длина белков — 484 и 442 аминокислоты. Большие белки (более 500 аминокислот) редко встречаются у бактерий, но составляют значительную часть (около 1/3) белков эукариот. Увеличение длины достигается за счет добавления дополнительных доменов, каждый из которых, как правило, состоит из 100–300 аминокислот. Вместе с тем увеличение размера белков очень слабо влияет на размер генома.

Подсчет активных генов позволяет взглянуть на их число с другой стороны. Если опираться на оценку числа различных видов мРНК, встречающихся в клетке, можно сделать вывод, что в среднем в клетке позвоночных экспрессируется от 10 000 до 20 000 генов. Значительное



РИС. 5.6 У мух *Drosophila* ~20% генов кодирует белки, участвующие в обслуживании генов (поддержании и экспрессии), и <10% генов кодирует белки, задействованные в обеспечении нормального хода клеточного цикла и различных сигнальных путей. Половина генов *Drosophila* кодирует белки, функция которых неизвестна

сходство популяций мРНК различных типов клеток позволяет предположить, что общее число активных генов (для данного организма) будет в несколько раз больше этого значения. Если принять во внимание, что геном человека содержит 20 000–25 000 генов (см. разд. 5.5 «В геноме человека меньше генов, чем кажется»), то следует признать, что в каждой отдельной клетке экспрессируется значительная часть от общего числа генов.

Эукариотические гены транскрибируются поодиночке. Каждый ген производит моноцистронную («одногенную») мРНК. Хотя из этого правила есть и исключения: в геноме *C. elegans* около 15% генов считываются в виде полицистронных пре-мРНК (видимо, для возможности использования *транс*-сплайсинга, позволяющего экспрессировать гены, расположенные существенно ниже промотора; см разд. 26.13 «В реакциях *транс*-сплайсинга участвуют малые РНК»).

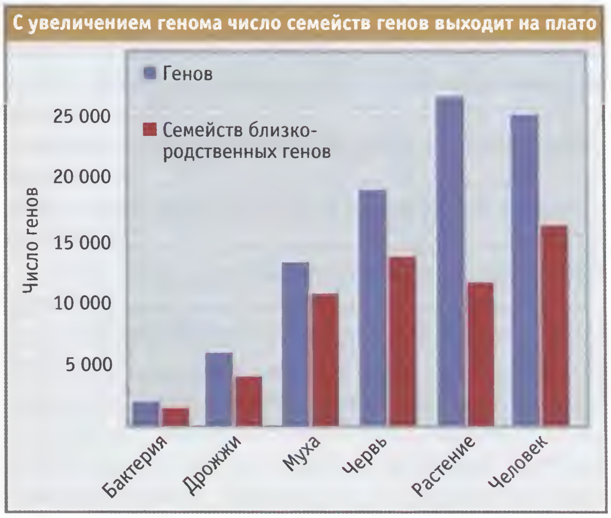


РИС. 5.7 В процессе эволюции многие гены дуплицируются; поэтому в геномах высших эукариот число различных семейств генов значительно меньше общего числа генов

5.4 Сколько существует различных типов генов?

Наряду с уникальными генами существуют множественные гены, образующие семейства, члены которых схожи (но обычно не идентичны) по последовательности. С увеличением размера генома доля уникальных генов уменьшается, а доля генов, принадлежащих семействам, увеличивается. У бактерий минимальное число семейств генов составляет >1000, у дрожжей — >4000, а у высших эукариот — от 11000 до 14000.

Некоторые гены представлены более чем одной копией (или очень схожи по последовательности с другими); таким образом, число разных типов генов меньше, чем общее число генов. Все гены можно разделить на группы, состоящие из схожих по последовательностям представителей, причем их сходство определяется сравнением экзонов. (Семейство генов возникает путем дупликации предкового гена и последующего накопления изменений в возникших копиях. Чаще всего члены семейства сходны, но не идентичны.) Общее число семейств генов можно определить путем прибавления к числу уникальных генов (для которых нет сходных по последовательности партнеров) к числу семейств, включающих два или более членов.

На **РИС. 5.7** приведено сравнение общего числа генов с числом различных семейств генов для шести эволюционно далеких геномов. Видно, что у бактерий подавляющее большинство генов уникальны, и число семейств близко по значению к общему числу генов. Однако уже у дрожжей *S. cerivisiae,* относящихся к низшим эукариотам, ситуация меняется, и доля повторяющихся генов у них возрастает. Интересно, что при переходе к высшим эукариотам число генов увеличивается скачкообразно, а число семейств генов практически не меняется.

На **РИС. 5.8** показано, что доля уникальных генов быстро уменьшается с увеличением размера генома.

Число семейств увеличивается вместе с размером генома

Организм	Уникальные гены	Семейства из 2–4 генов	Семейства из 4-х и более генов
H. influenzae	89%	10%	1%
S. cerevisiae	72%	19%	9%
D. melanogaster	72%	14%	14%
C. elegans	55%	20%	25%
A. thaliana	35%	24%	41%

РИС. 5.8 Доля генов, представленных в геноме несколькими схожими копиями, увеличивается вместе с размером генома и достигает максимума у высших эукариот

Если посмотреть на число представителей семейств, видно, что у бактерий и низших эукариот число членов для каждого семейства небольшое, а у высших эукариот — довольно высокое. Так, большая часть генов *Arabidopsis* входит в семейства, состоящие из 4 и более членов.

Если предположить, что экспрессируется каждый ген, то общее число генов должно примерно совпадать с общим числом белков, необходимым для формирования организма (протеомом). На самом деле две особенности влияют на то, что число белков, составляющих протеом, отличается от числа генов. Поскольку гены дуплицированы, некоторые из них кодируют один и тот же белок (хотя синтезирован он может быть в различное время и в разных местах), а некоторые кодируют похожие белки, которые опять же могут выполнять сходную функцию в разных местах и в разное время.

Кроме того, число белков, составляющих протеом, может превышать число генов, поскольку в результате альтернативного сплайсинга некоторые гены производят более одного белка.

Каков же минимальный размер протеома (набора различных типов белков, составляющих организм)? Оценить минимальное значение можно по числу семейств генов, увеличивающемуся в ряду организмов от 1400 у бактерий и несколько более 4000 у дрожжей до 11 000–14 000 у мух и червей.

Как распределен протеом по клетке? 6000 белков протеома дрожжей включают 5000 водорастворимых и 1000 мембранных белков. Около половины из них расположены в цитоплазме, четверть в ядре, и оставшаяся часть относится к митохондриям и эндоплазматическому ретикулуму или аппарату Гольджи.

Какое число генов является общим для всех организмов (или крупных групп типа бактерий или высших эукариот) и сколько генов специфичны для данного типа организмов? На РИС. 5.9 приведено сравнение генов дрожжевой клетки, червя и мухи. Гены, которые в различных организмах кодируют гомологичные белки, называют **ортологами**. Принято считать, что продукты двух генов в разных организмах могут выполнять сходные функции, если последовательности этих генов совпадают более чем на 80%. Руководствуясь этим критерием, следует предположить, что около 20% генов мух имеют ортологов у дрожжей. Скорее всего, эти гены необходимы всем эукариотам. При сравнении мухи и червя доля таких генов возрастает до 30%, что, по-видимому, означает добавление генов, общих для многоклеточных эукариот. Несмотря на это, основная часть генома по-прежнему представлена генами, кодирующими специфичные для мух или червей белки.



РИС. 5.9 В геноме *Drosophila* есть гены, которые, скорее всего, есть у всех эукариот. Дополнительная группа генов присутствует, скорее всего, у всех многоклеточных эукариот. Но все-таки большая часть генов данного генома специфична для представителей одного подрода

Создать представление о протеоме можно, исходя из количества и структуры генов, а можно — с помощью прямого исследования белкового содержимого клетки или всего организма. Эти подходы к исследованию белков позволили локализовать гены, которые сложно было определить на основании анализа последовательности генома. Для крупномасштабного исследования белков используют несколько методов. С целью идентификации белков, находящихся в смеси, полученной из клеток или тканей, применяют масс-спектрометрию. Другим методом является получение так называемых гибридных белков (несущих аффинные метки) путем экспрессии молекул кДНК с использованием генно-инженерных векторов. Такие кДНК-библиотеки позволяют одновременно анализировать множество белковых продуктов. Методы сравнения белков также эффективны при сравнении двух тканей (например, ткани здорового индивидуума и пациента, демонстрирующего то или иное заболевание) и установления существующих между ними отличий.

Теперь, когда нам известно общее число белков, мы можем попытаться выяснить, как они взаимодействуют друг с другом. По определению, белки, образующие структурные многокомпонентные комплексы, должны формировать стабильные связи друг с другом. Взаимодействие же белков сигнальных путей носит временный характер. И те, и другие типы взаимодействий можно выявить с помощью подходов, основанных на усилении эффекта исследуемого взаимодействия. Одной из наиболее популярных систем такого рода является двугибридный анализ, обсуждаемый в разд. 25.3 «Связывание ДНК и активация транскрипции осуществляются независимыми доменами фактора». Такой тип анализа не позволяет определить весь спектр взаимодействий в клетке: например, если в метаболическом пути один из ферментов высвобождает растворимый метаболит, который затем взаимодействует со следующим ферментом, то между участвующими в этом процессе белками прямые взаимодействия могут отсутствовать.

На практике попарный анализ взаимодействий между белками может дать нам информацию о минимальном числе независимых структур или путей. Тестирование всех 6000 (предсказанных) дрожжевых белков на взаимодействие в попарных комбинациях демонстрирует, что около 1000 белков могут связываться по крайней мере с одним партнером. Прямое исследование на формирование комплексов позволило определить 1440 различных белков в 232 многокомпонентных комплексах. Это лишь начальная стадия исследования, которое должно, в конечном счете, привести к определению числа функциональных комплексов и регуляторных путей. Сравнительный анализ 8100 белков человека выявил 2800 взаимосвязей, однако в контексте столь крупного протеома полученные результаты значительно сложнее интерпретировать.

Помимо функциональных генов существуют копии, утерявшие свою функцию (выявляемые по прерываниям кодирующей последовательности). Такие копии называют псевдогенами (см. разд. 6.6 «Псевдогены — тупики

эволюции»). Псевдогенов в геноме может быть много. В геномах мыши и человека на их долю приходится около 10% от числа (потенциально) активных генов (см. разд. 4.8 «Консервативность организации генома помогает выявить гены»).

С одной стороны, знать плотность генов нужно для того, чтобы оценить их общее число. Но одновременно следует задаться вопросом: имеет ли эта плотность самостоятельный биологический смысл? Существуют ли структурные ограничения, заставляющие гены быть разнесенными на определенные расстояния, и насколько сильно этот фактор обусловливает большие размеры геномов эукариот?



РИС. 5.10 Геном мыши содержит ~30 000 генов, кодирующих белки; этим генам соответствует ~4000 псевдогенов. В этом же геноме присутствует ~1600 генов, конечными продуктами которых являются различные не кодирующие белок РНК. Информация по этим генам представлена в правой части диаграммы, где масштаб крупнее. Сюда относятся ~800 генов рРНК, ~350 генов тРНК (которым соответствует ~150 псевдогенов) и ~450 генов других некодирующих белок РНК (например, малых ядерных РНК и микроРНК)

5.5 В геноме человека меньше генов, чем кажется

Ключевые понятия:

- Только 1% генома человека состоит из кодирующих регионов.
- Экзоны составляют около 5% от общей длины гена; сами гены (экзоны и интроны) составляют около 25% генома.
- Геном человека содержит от 20 000 до 25 000 генов.
- Около 60% генов человека могут экспрессироваться по механизму альтернативного сплайсинга.
- До 80% случаев альтернативного сплайсинга приводят к изменению последовательности белка; таким образом, протеом человека представлен 50 000–60 000 различных полипептидных цепей.

Среди позвоночных первым геномом, для которого была определена полная нуклеотидная последовательность, стал геном человека. Решение этой сложнейшей задачи дало огромный объем информации об устройстве нашего генома и его эволюции. Позднее, с определением последовательности генома мыши, стало возможным сравнение этих близкородственных геномов, что еще больше углубило наши знания.

По величине геномы приматов и грызунов практически равны: их размер составляет около 3×10^9 нуклеотидов (см. разд. 4.5 «Почему геномы такие большие?»). Мышиный геном примерно на 14% меньше человеческого, и вызвано это, вероятнее всего, повышенной частотой возникновения делеций. Подавляющая часть генов одного генома имеет ортологов в другом, то есть геномы содержат схожий набор генов и их семейств, однако по числу членов семейства могут различаться (в особенности это касается генов, кодирующих видоспецифичные функции) (см. разд. 4.8 «Консервативность организации генома помогает определять гены»). На сегодняшний день считается, что по числу генов геномы мыши и человека примерно равны (от 20 000 до 25 000 генов). На РИС. 5.10 приведена гистограмма распределения мышиных генов по группам. Из 30 000 генов, кодирующих белки, около 4000 являются псевдогенами. Примерно 800 генов отвечают за образование РНК, не кодирующих белки; в большинстве своем такие гены невелики (кроме рРНК) и примерно половина из них кодирует тРНК (для которых также выявлено большое число псевдогенов).

Гаплоидный геном человека состоит из 22 аутосом, а также X или Y-хромосомы. Длина хромосом варьирует от 45 до 279 Мпн и в общей сложности дает геном размером 3286 Мпн ($3,3 \times 10^9$ пн). На основании структуры хромосом геном можно разделить на эухроматин (потенциально содержащий активные гены) и гетерохроматин (см. разд. 28.7 «Хроматин подразделяют на эухроматин и гетерохроматин»). Эухроматин составляет большую часть генома — около $2,9 \times 10^9$ пн (т. е. примерно 90% последовательности). Помимо того, что последовательность нуклеотидов кодирует РНК и белки, она также определяет важные структурные особенности генома (см. разд. 28.8 «У хромосом проявляются полосатые паттерны окрашивания»).

На РИС. 5.11 показано, что экзоны (собственно белок-кодирующие последовательности) занимают лишь малую часть генома (около 1%). Интроны, составляющие большую часть гена, увеличивают общее количество вовлеченной в синтез белков ДНК до 25%. Как показано на РИС. 5.12, «средний» ген человека имеет длину в 27 тпн и состоит из 9 экзонов, в сумме образующих последовательность в 1340 пн. Таким образом, средняя белок-кодирующая последовательность составляет лишь 5% от общей длины гена.

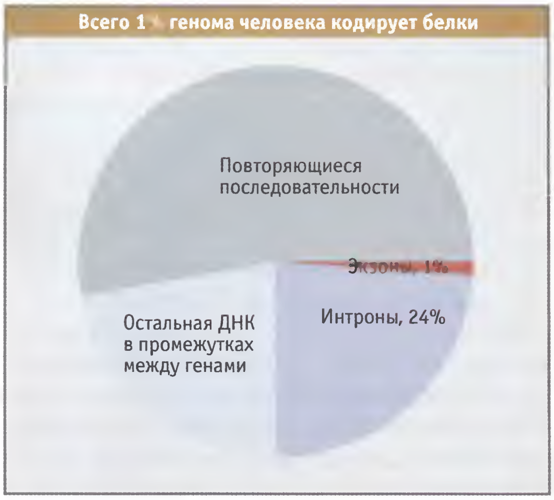


РИС. 5.11 **Хотя 25% генома человека занято генами, последовательности, которые на самом деле кодируют белки, составляют лишь крошечную часть этой фракции**

Две независимо предпринятые попытки определения полной последовательности генома человека дали оценку общего числа содержащихся в нем генов на уровне 30 000 и 40 000. Одним из способов оценки точности выполненного анализа последовательностей было сравнение выявленных генов. Каково же было удивление исследователей, когда перекрывание между двумя наборами генов составило лишь 50% (РИС. 5.13). Предыдущий анализ генов человека, основанный на РНК-транскриптах, обнаружил около 11 000 генов, причем почти все они присутствовали и в той, и в другой группе генов, определенных при анализе генома (находились в области перекрывания между этими группами), то есть достоверность выявления примерно половины генов человека почти не вызывала сомнений. Оставался вопрос со второй половиной каждого из вариантов трактовки генома. Наблюдаемые различия являлись следствием ошибок секвенирования столь крупного генома! При дальнейшем анализе последовательности (и по мере того, как определяли последовательности геномов, с которыми можно было производить сравнение) число достоверно локализованных генов уменьшалось, и по нынешним представлениям оно составляет примерно 20 000–25 000.

При любом современном способе оценки, общее число генов человека оказывается значительно меньше, чем считалось ранее. До того, как была определена полная последовательность генома человека, считалось, что в нем содержится порядка 100 000 генов. Геном человека демонстрирует лишь некоторое возрастание числа генов по сравнению с мухами и червями (13 600 и 18 500, соответственно), не говоря уже о геноме растения *Arabidopsis* (25 000) (рис. 5.2), хотя нас в принципе не должно удивлять то, что для построения более сложного организма нет необходимости в добавлении значительного числа генов. Например, разница между последовательностями ДНК человека и шимпанзе очень незначительна (менее 1%), показывая, что функции и взаимодействия схожих наборов генов могут приводить к различным результатам, а функции отдельных групп генов могут быть особенно значимыми. Подробное сравнение ортологов человека и шимпанзе демонстрирует, что определенные классы генов (включая гены, вовлеченные в процессы раннего развития, связанные с обонянием, слухом, и в целом с видоспецифичными функциями) претерпели значительные изменения.

Альтернативный сплайсинг является причиной того, что число генов меньше числа возможных белков, а он получил большее развитие у человека по сравнению с мухами и червями. Альтернативный сплайсинг у человека может затрагивать до 60% генов, увеличивая разницу между размерами протеома человека и других эукариот без пропорционального увеличения числа генов. Исследование генов, расположенных на двух хромосомах, показало, что вклад альтернативного сплайсинга в изменение последовательности белка может достигать 80%. Это может привести к увеличению числа белков, составляющих протеом, до 50 000–60 000.



РИС. 5.12 **Среднестатистический ген человека имеет длину 27 тпн. В нем девять экзонов: два длинных концевых экзона и семь внутренних экзонов. UTR, нетранслируемые (некодирующие) регионы, содержатся в каждом из концевых экзонов. (Все показатели рассчитаны как средние арифметические. У человека некоторые гены чрезвычайно длинные; с поправкой на них получаем среднюю длину 14 тпн с теми же семью экзонами.)**



РИС. 5.13 Два набора генов, идентифицированных в полностью секвенированном геноме человека, перекрываются лишь частично (как два больших круга вверху). Вместе оба набора включают в себя почти все ранее известные гены (это видно по перекрыванию больших кругов с меньшим кругом в нижней части диаграммы)

Следует отметить, что на уровне числа семейств генов отличия между человеком и другими эукариотами не так уж велики. Большинство генов человека входят в состав мультигенных семейств. В результате анализа около 25 000 генов, было обнаружено 3500 уникальных и 10 300 парных генов. Как видно из рис. 5.7, это лишь немногим больше, чем число семейств генов у червяка или мухи.

5.6 Как гены и прочие последовательности распределены по геному?

Ключевые понятия:

- Повторяющиеся последовательности (представленные более чем одной копией) составляют >50% генома человека.
- Значительный процент повторяющихся последовательностей представляет собой копии нефункциональных транспозонов.
- В геноме человека есть дупликации значительных по размеру регионов хромосом.

Равномерно ли распределены гены по геному? Некоторые хромосомы довольно бедны генами и более чем на 25% состоят из «пустынь» — участков последовательности более 500 тпн, не содержащих генов. Даже последовательности наиболее богатых генами хромосом более чем на 10% состоят из таких участков. В целом около 20% генома человека представлено «пустынными» участками, не несущими генов.

Как показано на **РИС. 5.14**, повторяющиеся последовательности составляют более 50% генома человека. Их делят на пять больших классов:

- Транспозоны (как активные, так и неактивные) составляют 45% генома. Все транспозоны представлены множественными копиями.
- Псевдогены (всего около 3000 — то есть 0,1% от всей ДНК). Это последовательности, возникающие в результате вставки ДНК-копии последовательности мРНК обратно в геном (см. разд. 6.6 «Псевдогены — это тупики эволюции»).
- Простые повторы (высокоповторяющаяся ДНК типа $(CA)_n$, составляющая около 3% генома).
- Сегментные дупликации (участки от 10 до 300 тпн, дуплицированные и перенесенные в новый регион), составляющие 5% генома. Причем лишь небольшая часть этих дупликаций расположена на одной и той же хромосоме, а большинство принадлежат разным.
- Тандемные повторы одной и той же последовательности (особенно часто располагающиеся в центромерах и на концах хромосом).

По последовательности генома человека можно судить о значимости транспозонов. Транспозоны могут самореплицироваться и вставляться в новые участки. При этом они могут функционировать как в качестве элементов ДНК (см. гл. 21 «Транспозоны»), так и в виде активной РНК (см. гл. 22 «Ретровирусы и ретропозоны»). Их распределение по геному человека представлено на рис. 22.18. Большая часть транспозонов генома человека нефункциональна; в активной форме существуют лишь немногие из них. То, что значительная часть генома занята этими последовательностями, означает, что они играли большую роль в формировании генома. Интересно, что некоторые из современных генов изначально возникли в виде транспозонов

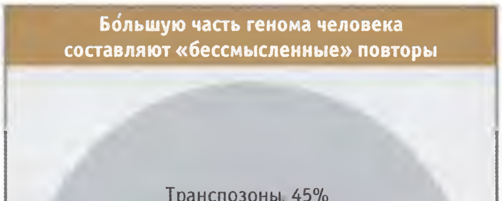


РИС. 5.14 Самый весомый компонент генома человека — это транспозоны. Кроме них есть и другие повторяющиеся последовательности (например, крупные дупликации и простые повторы)

и впоследствии эволюционировали в нынешнюю форму, потеряв при этом способность перемещаться по геному. Подобным образом возникло около 50 ныне существующих генов.

Сегментные дупликации в простом случае представляют собой тандемную дупликацию некоего региона в пределах одной хромосомы (как правило, по причине неправильной рекомбинации во время мейоза; см. разд. 6.7 «Неравный кроссинговер приводит к реорганизации генетических кластеров»). Однако во многих случаях дуплицированные регионы расположены на разных X-хромосомах, что свидетельствует либо о том, что они изначально были тандемной дупликацией, одна из копий которой затем была перенесена на новый участок, либо что дупликация по какому-то иному механизму сразу возникла на другой хромосоме. Крайним случаем сегментной дупликации является дупликация целого генома (например, когда изначально диплоидный геном превращается в тетраплоидный). Если дуплицированные копии накопят существенные различия, то геном может вновь стать диплоидным (хотя между копиями остается сходство, позволяющее судить о них как о результате дупликации). Такие события наиболее характерны для геномов растений. Анализ генома человека позволяет говорить о существовании множества отдельных дуплицированных регионов. Но установить, являются ли они результатом дупликации всего генома, произошедшей когда-то в линии позвоночных, сложно.

Любопытной особенностью генома человека является наличие последовательностей, демонстрирующих высокую эволюционную консервативность, но при этом не выполняющих белок-кодирующую функцию. По результатам сравнения с другими геномами (в первую очередь — с геномом мыши) такие последовательности составляют около 5% от всего генома. Возможно, эти последовательности являются кодирующими с функциональной точки зрения? Например, плотность расположения таких участков на 18-й хромосоме такая же, как и на любой другой, хотя концентрация белок-кодирующих генов на ней значительно ниже. Это отчасти подтверждает, что функция данных регионов не связана со структурой или экспрессией генов, кодирующих белки.

5.7 В Y-хромосоме есть несколько генов, специфичных для самцов

Ключевые понятия:

- Y-хромосома содержит около 60 генов, экспрессирующихся только в семенниках.
- Специфичные для самцов гены представлены множественными копиями в повторяющихся сегментах Y-хромосомы.
- Генная конверсия, происходившая между множественными копиями генов Y-хромосомы, позволила активным генам сохраниться в эволюции.

Определение последовательности человеческого генома существенно продвинуло нас в понимании роли половых X-хромосом. Считается, что X- и Y-хромосомы произошли от одной очень древней аутосомы. В процессе эволюции X-хромосома сохранила большинство исходных генов, а Y-хромосома практически все утеряла.

X-хромосома ведет себя как аутосома, поскольку у женщин она представлена в двух копиях — между ними может происходить рекомбинация. Плотность расположения генов на X-хромосоме сравнима с таковой у других хромосом.

Y-хромосома по размеру гораздо меньше, чем X, и содержит значительно меньше генов. Особенность Y заключается в том, что она встречается только у мужчин и в единственной копии, то есть локусы, сцепленные с Y-хромосомой, не диплоидны, как все остальные гены человека, а гаплоидны.

На протяжении многих лет считалось, что Y-хромосома несет исключительно гены, определяющие развитие мужского организма. Действительно, большая часть Y-хромосомы (>95% ее последовательности) не способна к кроссинговеру с X-хромосомой, на основании чего возникло представление об отсутствии в Y-хромосоме активных генов (поскольку кроссинговер является одним из способов предотвращения накопления делеций). Однако этот регион фланкирован короткими псевдоаутосомными участками, способными во время мейоза (у самцов) обмениваться информацией с X-хромосомой. Регион Y-хромосомы, расположенный между короткими псевдоаутосомными участками был изначально назван нерекомбинирующим, однако теперь он называется самецспецифичным регионом.

Детальный анализ последовательности Y-хромосомы показал, что самецспецифичный регион состоит из трех типов участков (РИС. 5.15):

- *Последовательности, перенесенные из X-хромосомы* состоят из 3,4 Мпн, собранных в крупные блоки и возникших в результате их переноса в Y-хромосому из локуса q21 X-хромосомы около 3–4 млн лет назад. Эти последовательности специфичны для людей, они не рекомбинируют с X-хромосомой и большей частью неактивны. На сегодняшний день известно только два активных гена в составе таких последовательностей.
- *X-вырожденные сегменты* Y-хромосомы несут последовательности, имеющие общее происхождение с X-хромосомой (возвращаясь к общей аутосоме, от которой произошли X- и Y-хромосомы), и содержат гены (или псевдогены), гомологичные генам X-хромосомы. Сюда относят 14 активных генов и 13 псевдогенов. Причем активные гены каким-то образом противостоят давлению отбора, обычно приводящему к исчезновению активных генов из не рекомбинирующих в мейозе регионов хромосомы.
- *Ампликонные сегменты* состоят из 10,2 Мпн и многократно повторяются на Y-хромосоме. Су-



РИС. 5.15 Y-хромосома включает в себя вырожденные X-сегменты, транслоцированные X-сегменты, а также области-ампликоны. Вырожденные и транслоцированные X-сегменты содержат, соответственно, два и четырнадцать генов; каждый из этих генов представлен единственной копией. В ампликонах содержится восемь крупных палиндромов (P1–P8) с генами девяти семейств. Каждое из этих семейств представлено как минимум двумя копиями

ществует восемь больших палиндромных блоков. Они включают девять семейств генов (кодирующих белки) по 2–35 копий на семейство. Название «ампликонные» отражает тот факт, что эти последовательности были внутренне амплифицированы на Y-хромосоме.

В целом в перечисленных регионах Y-хромосомы содержится больше генов, чем ожидалось: 156 транскрипционных единиц, половина из которых представляет гены, кодирующие белки, а половина — псевдогены.

Сохранение в Y-хромосоме активных генов, по всей видимости, связано с тем, что присутствие нескольких гомологичных копий генов в ампликонных сегментах позволяет использовать переход генов между множественными копиями для восстановления активных копий. Обычно необходимость во множественных копиях гена носит количественный (нужно много белка) или качественный (нужно несколько похожих или синтезируемых в разное время и в разных местах белков) характер. В случае же Y-хромосомы дополнительной функцией множественных копий генов является эволюционная функция. Существование множественных копий позволяет осуществлять рекомбинацию в пределах самой Y-хромосомы. Это необходимо для создания эволюционного разнообразия, в норме формируемого за счет рекомбинации между гомологичными хромосомами.

Большинство генов, расположенных в ампликонных сегментах, экспрессируется в семенниках и, вероятно, вовлечено в процесс формирования мужского организма. Если из общего числа генов (около 25 000) «мужских» генов около 60, то можно сказать, что генетические различия между мужчинами и женщинами составляют всего 0,2%.

5.8 Усложнение организмов в ходе эволюции происходило за счет появления у генов новых функций

Ключевые понятия:

- Сравнение различных геномов демонстрирует достоверное увеличение числа генов с постоянным появлением новых генов в ряду низших эукариот, многоклеточных организмов, животных и позвоночных.
- Большинство генов, встречающихся только у позвоночных, относятся к иммунной или нервной системе.

Сравнение последовательности генома человека с последовательностями других видов позволяет узнать много нового о процессе эволюции. На РИС. 5.16 приведен анализ генов человека с точки зрения их распространения в природе. 21% генов человека является общим для эукариот и прокариот (верхний правый угол рисунка). Они в большинстве своем кодируют белки, необходимые всем живым организмам: участвующие в базовых метаболических процессах, репликации, транскрипции и трансляции. Двигаясь по часовой стрелке, мы обнаруживаем, что 32% генов встречаются у всех эукариот (их можно найти даже у дрожжей). Эти гены кодируют белки, вовлеченные в основные функции эукариотических клеток (но не бактерий), — например, они могут быть связаны с органеллами или компонентами цитоскелета. Следующие 24% генов характерны для всех животных. Они включают гены, необходимые для формирования многоклеточности и для развития различных типов тканей. 22% генов встречаются только у позвоночных. В большинстве своем они кодируют белки иммунной и нервной систем. Вместе с тем они кодируют лишь не-



РИС. 5.16 Гены человека можно классифицировать на основании широты распространения их гомологов у представителей разных групп организмов



РИС. 5.17 У всех эукариот ключевыми клеточными функциями заведуют примерно одни и те же белки

большое число ферментов, что свидетельствует о древнем происхождении ферментативных систем и о том, что метаболические пути возникли на ранних стадиях эволюционного процесса. Таким образом, мы видим, что при движении от бактерий к позвоночным возникает необходимость в добавлении групп генов, кодирующих новые функции, необходимые на каждой новой ступени эволюции.

Общие для всех протеомов полипептиды, скорее всего, будут относиться к наиболее ключевым белкам живых систем. При детальном сравнении протеома человека с протеомами других организмов можно заметить, что 46% протеома дрожжей, 43% протеома червя и 61% протеома мухи представлены в протеоме человека. Присутствующий во всех четырех протеомах «базовый набор» состоит примерно из 1300 белков. Большая часть общих белков относится к так называемой группе белков «домашнего хозяйства» (housekeeping), отвечающих за ключевые функции и разделяемых на несколько типов, представленных на РИС. 5.17. Основные их функции связаны с транскрипцией и трансляцией (35%), метаболизмом (22%), транспортом веществ (12%), репликацией и модификацией ДНК (10%), укладкой и деградацией белков (8%), и общеклеточными процессами (6%).

В сравнении с протеомами других эукариот, протеом человека приобрел множество новых белков, но число белковых доменов при этом практически не изменилось. Большая часть белковых доменов является общей для представителей царства животных. Новые белки формируются в основном за счет комбинирования существующих доменов. На РИС. 5.18 показано, что для высших эукариот наибольший прирост числа новых белков происходит в группах трансмембранных и внеклеточных молекул. У дрожжей подавляющее большинство новых структур связано с внутриклеточными белками. У мухи (как и у червя) разнообразие внутриклеточных белков увеличивается вдвое, но при этом (на фоне возникающей необходимости сложных взаимодействий между клетками многоклеточного организма) значительно увеличивается число трансмембранных и внеклеточных белков. У позвоночных (человека) число внутриклеточных структур увеличивается незначительно, но при этом вновь

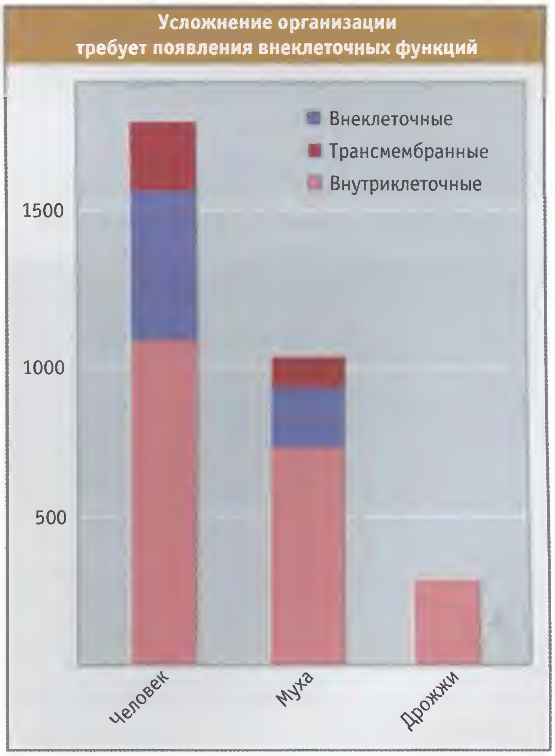


РИС. 5.18 Увеличение сложности организации эукариотических организмов сопровождается накоплением разновидностей белков, обслуживающих трансмембранные и внеклеточные функции

происходит сильное увеличение разнообразия трансмембранных и внеклеточных молекул.

Известно, что генетические различия между человеком и шимпанзе (нашим ближайшим родственником) невелики. Наши геномы совпадают примерно на 99%. Определение полной последовательности генома шимпанзе позволяет нам более детально исследовать 1% отличий и выяснить, можно ли среди этих отличий найти особенности, ответственные за формирование присущих человеку признаков. Сравнение геномов позволяет выявить 35×10^6 однонуклеотидных замен (всего около 1,2% различий между последовательностями), 5×10^6 делеций или инсерций (приводящих к формированию примерно 1,5% специфичных для каждого вида эухроматических последовательностей) и множество хромосомных перестановок. Гомологичные белки человека и шимпанзе, как правило, очень похожи; 29% полипептидов вообще идентичны, а для большинства остальных молекул различия между видами представлены одной или двумя аминокислотными заменами на белок. Интересно, что в генах, кодирующих белки, определяющие характерные для человека признаки, нуклеотидные замены встречаются даже реже, чем в среднем по генам. Поэтому можно предположить, что эволюция белков не играла главную роль в формировании различий между человеком и шимпанзе. В таком случае основными кандидатами на роль «человекообразующих» становятся более масштабные изменения в структуре генома и/или изменения механизмов регуляции работы генов. Заметим, что примерно 25% однонуклеотидных замен затрагивают CpG динуклеотиды (большинство из которых являются потенциальными регуляторными сайтами).

5.9 Какая часть генов действительно необходима?

Ключевые понятия:

- Не все гены одинаково важны. У дрожжей и мух удаление единственного гена из генома приводит к видимому эффекту лишь для каждого второго гена.
- Если два или более генов являются продублированными (вырожденными), то мутация в одном из них может не оказывать видимого эффекта на фенотип.
- Мы не до конца понимаем причины существования в геноме явно необязательных генов.

Естественный отбор — это сила, гарантирующая сохранение в геноме полезных генов. Возникновение мутаций носит случайный характер, и при их попадании в открытую рамку считывания наиболее вероятным эффектом станет нарушение структуры белка. Организм, несущий повреждающую белок мутацию, с эволюционной точки зрения окажется в невыгодном положении, и в конце концов носитель такой мутации (обладающий пониженной жизнеспособностью) будет элиминирован. Частота аллеля в популяции определяется балансом между процессами формирования новых аллелей и элиминированием старых. Если взглянуть на это правило с другой стороны, можно сказать, что при обнаружении консервативной открытой рамки считывания следует допустить, что ее продукт играет важную роль в организме. Естественный отбор противодействует накоплению изменений в гене. Вместе с тем «необязательные» гены могут быстро накапливать мутации, иногда изменяя исходную последовательность до неузнаваемости.

Эволюционная консервативность гена означает, что он дает организмам селективное преимущество. Причем в процессе эволюции даже малейшее преимущество становится предметом естественного отбора (однако мутации далеко не всегда сказываются на фенотипе). Интересно узнать, сколько генов в действительности *необходимо* живой системе (в том смысле, что отсутствие любого из них летально для организма)? Для диплоидных организмов признаком необходимости гена будет летальность гомозиготной нуль-мутации.

Можно предположить, что доля необходимых генов будет падать с увеличением размера генома (поскольку более крупные геномы могут содержать множество схожих по функциям копий определенного гена). Однако это допущение не вполне согласуется с имеющимися данными (см. рис. 5.2).

Одним из подходов, используемых для определения числа необходимых генов, является мутационный анализ. Если мы насыщаем летальными мутациями некий участок хромосомы, то на карте эти мутации должны попасть в определенное число комплементационных групп, соответствующее числу летальных локусов в этом регионе. Перенося эту ситуацию на весь геном, можно посчитать, сколько всего необходимых генов в нем содержится.

У организма с самым маленьким (из известных) геномом (*M.genitalium*) случайные инсерции приводят к заметному эффекту лишь у двух третей генов. С похожей ситуацией мы сталкиваемся и в случае с *E.coli*, у которой необходимыми являются менее половины генов. У дрожжей *S.cerevisiae* доля необходимых генов еще меньше. Анализ результатов случайных вставок в их геном показал, что лишь 12% инсерций обладали летальным эффектом, а еще 14% подавляли рост. Большинство (70%) инсерций вообще не оказывали видимого эффекта. Более глубокое исследование, заключавшееся в поочередном удалении каждого из 5916 генов дрожжей (>96% известных генов), показало, что только 18,7% из них важны при росте на богатых питательными веществами средах. На РИС. 5.19 показано, что в состав «важнейших» генов попали гены практически всех категорий. Самая значительная часть этих генов (около 50%) кодирует продукты, вовлеченные в синтез белка. Не будем забывать, что подход с отбором мутантов на богатой среде не позволяет выявить все гены, ответственные за жизнедеятельность дрожжей в условиях естественной среды (где набор питательных веществ, как правило, ограничен).

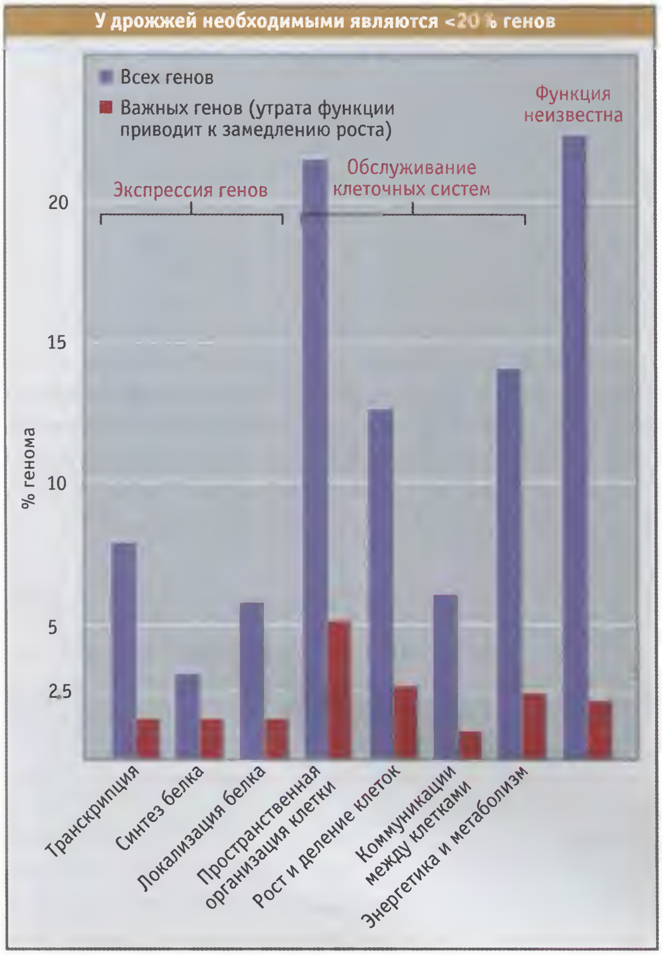


РИС. 5.19 Во всех функциональных категориях генов *S. cerevisiae* есть необходимые гены

РИС. 5.20 суммирует результаты анализа эффектов от систематического удаления генов червя *C. elegans*. Регионы генов были предсказаны путем компьютерного анализа полной последовательности генома, а их выключение проводили при помощи прицельного уничтожения транскриптов методом интерферирующих РНК (см. разд. 13.10 «РНК-интерференция — причина молчания»). Таким образом были получены черви, в каждом из которых функция одного из предсказанных генов была подавлена. Только у 10% этих «нокаутов» были обнаружены эффекты, оказывающие видимое влияние на фенотип, а, следовательно, большая часть генов червей не играет исключительной роли. Гораздо большая доля значимых генов (21%) была показана для локусов, имеющих гомологов у других эукариот (то есть консервативные гены в большей степени связаны с ключевыми функциями организма). Кроме того, доля «необходимых» генов оказалась выше и для последовательностей, существующих в единственной копии на гаплоидный геном (в сравнении с множественными копиями гомологичных или идентичных генов). Можно предположить, что значительная часть множественных генов червей представляет собой относительно недавние дупликации, способные заменять друг друга в выполнении тех или иных функций.

Подробный анализ числа значимых генов у высших эукариот был проведен и для *Drosophila*. Суть анализа заключалась в попытке найти взаимосвязь между видимыми (в микроскоп) особенностями хромосомной структуры и числом функциональных генетических единиц. Идея о возможности такого анализа возникла благодаря наличию видимых глазом полос в политенных X-хромосомах *D. melanogaster*. (Эти хромосомы встречаются на определенных стадиях развития и представляют собой мультиплицированные хромосомы, на которых хорошо заметно (в микроскоп) чередование компактизованных (неактивных) и декомпактизованных (активных) регионов; см. разд. 28.10 «Политенные хромосомы выглядят полосатыми».) Предположив, что эти полосы представляют собой гены, можно попробовать определить корреляцию между организацией генов и организацией полос. Гаплоидный набор хромосом *D. melanogaster* формирует около 5000 полос; эти полосы отличаются по размеру примерно в 10 раз, однако в среднем их длина составляет около 20 тпн на полосу.

Простейшим подходом к локализации гена является насыщение региона хромосомы мутациями. Обычно регистрируют летальные мутации без определения причины смерти. Любая летальная мутация служит индикатором крайне значимого для организма локуса. Иногда мутации приводят к формированию видимых нарушений, и в таком случае, даже если они не приводят к летальному исходу, их тоже можно использовать в качестве индикатора значимого локуса. После распределения мутаций по комплементационным группам число этих групп можно сравнить с числом полос в данном регионе. Отдельные комплементационные группы даже можно приписать к определенным полосам. В результате этих экспериментов исследователи пытались установить, есть ли четкая взаи-



РИС. 5.20 Систематический анализ утраты функции, проведенный для 86% генов *C. elegans*, свидетельствует о том, что нарушение работы лишь 10% генов проявляется фенотипически

мосвязь между полосами и генами, например справедливо ли то, что в каждой полосе находится один ген?

Обобщая анализ, проводимый в течение последних 30 лет, можно сказать, что число летальных комплементационных групп составляет около 70% от числа полос. Открытым остается вопрос, существует ли функциональная взаимосвязь генов и полос. Оценка числа летальных генов *Drosophila* дает значение порядка 3600. При любом способе измерения, число летальных локусов у *Drosophila* значительно меньше общего числа генов.

Если доля «незаменимых» генов человека схожа с таковой у других эукариот, можно предсказать наличие 4000–8000 генов, мутации в которых будут летальными или будут приводить к очевидным негативным эффектам. На сегодняшний день выявлено около 1300 генов человека, мутации в которых вызывают видимые дефекты. Учитывая тот факт, что летальные мутации, как правило, действуют на столь ранних стадиях развития, что увидеть их эффект практически невозможно, число таких мутаций может составлять весомую долю от общего числа летальных или приводящих к фенотипическим дефектам мутаций. Этот эффект, в частности, может объяснить результаты, представленные на РИС. 5.21, где показано, что большинство известных генетических дефектов человека вызвано точечными мутациями (ген, как правило, сохраняет ослабленную функцию).

Как же объяснить наличие в геноме генов, удаление которых не приводит к видимому эффекту? Вероятнее всего, в организме существуют альтернативные пути реализации одной и той же функции. Такая функциональная **избыточность** возможна благодаря существованию множественных копий генов. Подтверждение этому мы находим в экспериментах, в которых для получения эффекта необходимо «выключить» все копии данного гена. Более того, организм может обладать и вовсе независимыми путями выполнения определенной функции. Инактивация одного из продублированных путей, скорее всего, не даст фенотипического эффекта, однако одновременное нарушение всех путей приведет к нарушению функции и фенотипа.

Большинство генетических дефектов человека обусловлено небольшими мутациями	
Миссенс и нонсенс	58%
Ошибки сплайсинга	10%
Мутации в регуляторных областях	<1%
Небольшие делеции	16%
Небольшие инсерции	6%
Крупные делеции	5%
Крупные перестройки	2%

РИС. 5.21 Большинство известных наследственных заболеваний человека обусловлено точечными мутациями. Большая часть этих мутаций напрямую влияет на аминокислотную последовательность конкретных белков. Остальные генетические дефекты обусловлены разных размеров вставками, делециями и перестройками

Случаи дублирования функции можно выявить, комбинируя мутации. Например, в два гена одного и того же штамма вводят делеции, по отдельности не являющиеся летальными. Если двойной мутант оказывается нежизнеспособным, штамм называют **синтетически летальным**. Разработка методов эффективного обнаружения двойных мутантов позволила достичь значительных успехов в изучении дрожжей. Метод получил название **SGA** (генетический анализ на чипах, synthetic genetic array analysis). На РИС. 5.22 представлены результаты SGA, проведенной для каждого из 132 жизнеспособных мутантов путем тестирования на выживаемость в комбинации с каждой из 4700 жизнеспособных делеций. Каждый из тестируемых генов дал, по крайней мере, одного партнера, комбинация с которым была летальной. Большинство же генов обнаружило множество таких партнеров (в среднем — около 25). Для одного из генов было обнаружено наибольшее число летальных партнеров — 146. Небольшая доля взаимодействующих мутантных пар (около 10%) кодировала физически взаимодействующие белки.

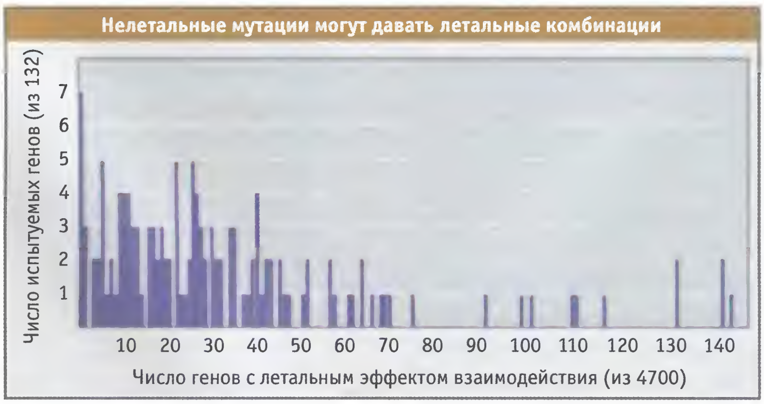


РИС. 5.22 132 мутантных испытуемых гена дают летальные комбинации с нелетальными мутациями в других 4700 генах. Диаграмма отражает количество летальных сочетаний, полученных для каждого из испытуемых генов

Полученные на дрожжах результаты позволяют продвинуться в объяснении слабости эффекта большинства случайных делеций. Естественный отбор действует против таких делеций, только если они присутствуют в попарных летальных комбинациях. В некоторой степени организм защищает себя от разрушительного воздействия мутаций путем формирования функциональной избыточности. Однако платой за это становится повышение «генетической нагрузки», заключающейся в накоплении мутаций, не вызывающих потери функции, но способных вызывать серьезные проблемы в комбинации с другими мутациями в последующих поколениях. Исходя из теории естественного отбора, можно предположить, что утрата некоторых генов способствует сохранению других генов в ходе эволюции.

5.10 Гены отличаются по степени экспрессии

Ключевые понятия:

- В клетке большинство генов экспрессируется на сравнительно низком уровне
- Только небольшое число генов, определяющих данный тип клеток, экспрессируются на высоком уровне.

По количеству ДНК, способной гибридизоваться с РНК, можно определить долю ДНК, представленную в мРНК. Такой анализ обычно выявляет около 1% ДНК (очевидно, способной служить матрицей для синтеза мРНК). Основываясь на этом факте и исходя из знания средней длины мРНК, можно примерно определить общее число работающих в данный момент генов. Для низших эукариот (например, для дрожжей) число одновременно работающих генов обычно составляет около 4000. В соматических тканях высших эукариот это число колеблется от 10 000 до 15 000. Примерно такие же значения отмечены для растений и позвоночных. (Важным исключением является мозг млекопитающих, в котором работает значительно больше генов, однако точное их число неизвестно.)

Для определения структуры популяции мРНК можно использовать анализ кинетики реассоциации. Обычно этот тип анализа позволяет разбить весь пул мРНК эукариотической клетки на 3 составляющих субпопуляции (РНК-компонента). Как и в случае с кривой реассоциации ДНК, каждый из компонентов реассоциирует в интервале, захватывающем примерно два порядка значений показателя R_0t (величины, отражающей молекулярную концентрацию РНК, помноженную на время реакции). Разброс значений R_0t, выходящий за пределы двух порядков величины, относят на счет наличия в пуле других составляющих. Описанные «концентрационные максимумы» наиболее эффективно выявляются при помощи математического сглаживания суммарной кривой компьютерными методами. Тем не менее очевидно, что кривая реассоциации отражает непрерывный спектр нуклеотидных последовательностей в клеточном пуле мРНК.

Пример трехкомпонентной кривой, характеризующей гибридизацию избытка мРНК с кДНК, приведен на **РИС. 5.23**:

- Первый компонент реакции по всем характеристикам соответствует взаимодействию овальбуминовой мРНК со своей кДНК в контрольной реакции. Это означает, что первый компонент, по сути, является мРНК овальбумина (которая на самом деле занимает около половины массы мРНК в ткани яйцевода).
- Следующий компонент реакции включает около 15% молекул с общей сложностью 15 тпн, что соответствует 7–8 видам мРНК со средней длиной в 2000 оснований.
- Последний компонент включает 35% молекул общей сложностью 26 Мпн. Это соответствует приблизительно 13 000 видов мРНК со средней длиной в 2000 оснований.

Из приведенных данных можно сделать вывод, что около половины от общей массы мРНК в клетке представлено единственным видом мРНК, 15% массы представлено 7–8 видами мРНК, и 35% массы разделено между 13 000 видов мРНК. Следовательно, представляющие каждый компонент молекулы мРНК должны присутствовать в клетке в разных количествах.

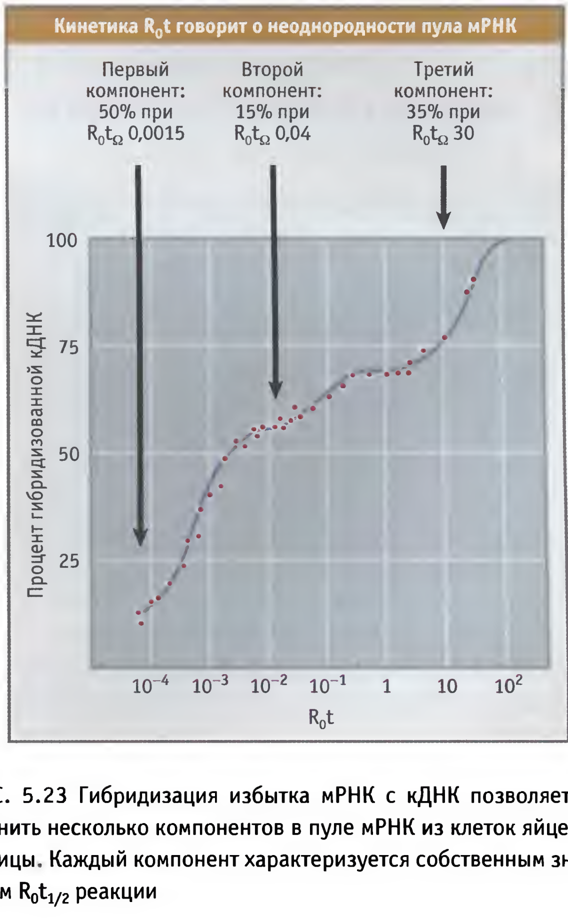


РИС. 5.23 Гибридизация избытка мРНК с кДНК позволяет вычленить несколько компонентов в пуле мРНК из клеток яйцевода курицы. Каждый компонент характеризуется собственным значением $R_0t_{1/2}$ реакции

Среднее число молекул мРНК каждого типа на клетку называют их **представленностью**. Если известно общее число молекул мРНК на клетку, то представленность каждого класса молекул можно легко посчитать. В примере, приведенном на рис. 5.23, суммарная мРНК включает 100 000 копий первого компонента (мРНК овальбумина), около 4000 копий каждого из 7–8 видов мРНК второго компонента и в среднем лишь по 5 копий каждой из 13 000 мРНК, составляющих третий компонент.

По представленности популяцию мРНК можно разделить на два основных класса:

- Отметим, что так много (как в приведенном примере) мРНК одного единственного гена встречается довольно редко (в примере это мРНК овальбумина в клетках яйцевода), в большинстве же случаев в клетке содержится несколько разных **высокопредставленных мРНК**. Этот компонент обычно состоит из <100 различных мРНК, представленных 1000–10 000 копий на клетку. Как правило, этот компонент составляет значительную долю (порядка 50%) от всей мРНК клетки.
- Оставшаяся половина мРНК состоит из большого числа разных по последовательности молекул (порядка 10 000), каждая из которых представлена лишь небольшим числом копий (менее 10). Этот класс называют **низкопредставленной (редкой),** или **комплексной, мРНК.** Именно этот класс определяет выход графика на плато.

5.11 Сколько генов экспрессируется в каждый момент времени?

Ключевые понятия:

- Наборы мРНК, производящиеся в клетке на низком уровне, очень похожи для различных типов клеток.
- мРНК, представленные в клетке в большом количестве, часто специфичны для данного типа клеток.
- У высших эукариот около 10 000 разных транскриптов могут быть обнаружены в большинстве типов клеток организма.

В большей части соматических тканей высших эукариот работает от 10 000 до 20 000 генов. Каково же перекрывание наборов генов, работающих в различных тканях? В печени курицы экспрессируется примерно от 11 000 до 17 000 генов, а в яйцеводе — 13 000–15 000. Какая доля транскриптов в этих тканях общая, а какая — тканеспецифична? Ответ на эти вопросы можно получить, анализируя транскриптом — всю мРНК данной клетки или ткани.

Как правило, наибольший вклад в формирование различий между транскриптомами вносят мРНК, относящиеся к классу преобладающих. Например, овальбумин синтезируется только в яйцеводе и вовсе не обнаруживается в печени. Это означает, что как минимум половина от всей мРНК яйцевода специфична для данного типа ткани.

Однако высокопредставленные мРНК составляют лишь малую часть от общего числа работающих генов. Для понимания транскрипционных различий, существующих между различными типами клеток, гораздо важнее определить разницу в составе фракции низкопредставленной мРНК.

Сравнение транскриптомов различных тканей показывает, что, например, около 75% мРНК для печени и яйцевода общие. Другими словами, около 12000 генов экспрессируется и в печени, и в яйцеводе, а специфичными для этих тканей являются около 5000 и 3000 генов соответственно.

Низкопредставленная мРНК для разных тканей сильно перекрывается. В клетках печени и почки мыши совпадает около 90% низкопредставленных мРНК, что оставляет на различия между этими тканями (выраженное в числе активных генов) 1000–2000 генов. Общий результат, полученный при нескольких сравнениях подобного рода, состоит в том, что лишь 10% от числа разных мРНК клетки тканеспецифичны. Большинство транскриптов являются общими для большинства (возможно, и для всех) типов клеток.

Это означает, что число генов, ответственных за выполнение ключевых функций (необходимых всем типам клеток), составляет у млекопитающих порядка 10 000. Иногда гены, выполняющие такие функции, называют **генами домашнего хозяйства** (housekeeping genes), или **конститутивными генами** (constitutive genes). Их противопоставляют генам, выполняющим специализированные функции (как, например, гены овальбумина или глобина), необходимые только клеткам определенного типа. Такие гены иногда называют **генами роскоши** (luxury).

5.12 Число экспрессирующихся генов можно определить за один подход

Ключевые понятия:

- Технология «чипов» позволяет быстро «сфотографировать» экспрессию всего генома дрожжевой клетки.
- Примерно 75% (около 4500 генов) генома дрожжей экспрессируется в нормальных условиях.
- С помощью чипов можно провести подробное сравнение для разных клеток животных, например, с целью определения различий в экспрессии генов между нормальной и опухолевой клетками.

Современные технологии позволяют осуществлять более глубокую и точную оценку числа работающих в данный момент генов. Первый подход (SAGE, serial analysis of gene expression) предполагает использование уникальной маркерной последовательности для

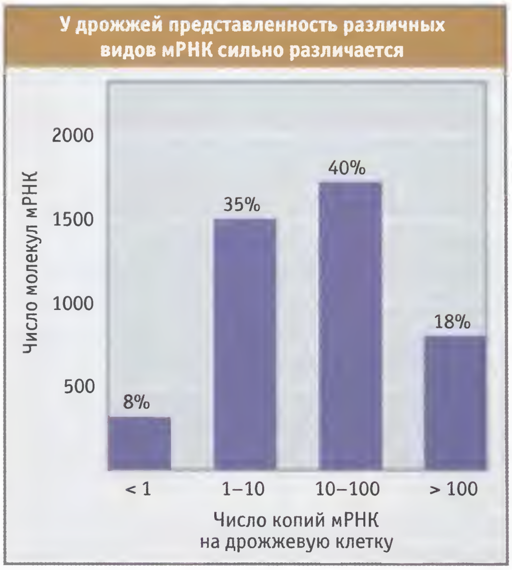


РИС. 5.24 У дрожжей весь спектр различных видов мРНК разбросан в интервале от менее одной на клетку (это означает, что такая мРНК есть не в каждой клетке) до более 100 (такие мРНК кодируют самые востребованные белки)

каждой мРНК. Технология позволяет определить относительное содержание каждого такого маркера в смеси. Этот подход выявляет 4665 работающих генов у растущих в нормальных условиях *S.cerivisiae* с относительным числом транскриптов каждого типа от 0,3 до свыше 200 на клетку. Это означает, что в данных условиях экспрессируется около 75% от общего числа генов (примерно 6000 для дрожжей). На **РИС. 5.24** приведено количество отдельных мРНК для каждого из уровней представленности.

Одна из наиболее мощных технологий оценки транскрипционного профиля использует так называемые чипы: набор присоединенных к некоторой поверхности олигонуклеотидов с заранее известной последовательностью (high-density oligonucleotide arrays, HDA). В результате определения полных последовательностей геномов стало возможным создание чипов, содержащих геноспецифичные олигонуклеотиды для каждого гена данного организма. В случае с *S. cerivisiae* каждая из 1681 открытых рамок считывания представлена на чипе 20 олигонуклеотидами, каждый их которых состоит из 25 нуклеотидов, в точности соответствующих последовательности мРНК, и 20 олигонуклеотидами, отличающимися от транскрипта на один нуклеотид. Уровень экспрессии каждого гена высчитывается путем вычитания фонового сигнала (полученного от частично несовпадающего олигонуклеотида) из сигнала, полученного от полностью комплементарного олигонуклеотида. Весь дрожжевой геном можно «разместить» всего на четырех таких чипах. Причем технология достаточно чувствительна для того, чтобы детектировать транскрипты 5460 генов (это около 90% от всех генов). Оказалось, что многие гены экспрессируются на очень низком уровне, с представленностью в 0,1–0,2 транскрипта на клетку (представленность менее 1 транскрипта на клетку означает, что в данный момент не все клетки содержат данный транскрипт).

Технология позволяет не только измерять уровень экспрессии генов, но и выявлять различия в профилях экспрессии в различных системах: мутантных клетках по сравнению с клетками дикого типа, между клетками, растущими в различных условиях и т. д. Результаты сравнения двух состояний представляются в виде сетки, в которой каждая клетка представляет определенный ген, а цвет выражает относительные изменения уровня его экспрессии в изучаемой системе. Левая часть **РИС. 5.25** демонстрирует эффект от мутации в гене РНК-полимеразы II (фермента, синтезирующего мРНК), которая, как можно предположить, значительно уменьшает уровень экспрессии большинства генов. Правая часть показывает, что мутация во вспомогательном компоненте аппарата транскрипции (*SRB10*) оказывает более слабый эффект и, в некоторых случаях, напротив, вызывает повышение уровня экспрессии генов.

Перенос этой технологии на животную клетку позволяет превратить довольно общие сведения, полученные на основании гибридизации отдельных РНК, в более точные данные о том, какие именно гены экспрессируются и каков уровень представленности их транскриптов в данном типе клеток. На определенной стадии жизненного цикла, карта генной экспрессии *D.melanogaster* демонстрирует транскрипционную активность практически всех (93%) предсказанных генов и показывает, что 40% из них представлены формами, полученными в результате альтернативного сплайсинга.

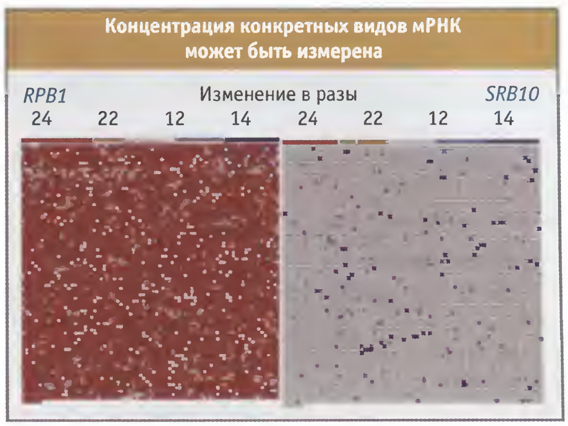


РИС. 5.25 Современные технологии позволяют фиксировать изменение в экспрессии каждого гена. Каждая крохотная клеточка представляет один ген (верхняя левая соответствует первому гену хромосомы I, нижняя правая — последнему гену хромосомы XVI). Изменение в экспрессии по отношению к дикому типу обозначено цветами: красным (уменьшение), синим (увеличение) или белым (без изменений). Фотография любезно предоставлена Rick A. Young, Whitehead Institute, Massachusetts Institute of Technology

5.13 Резюме

На сегодняшний день установлена полная последовательность геномов многих бактерий и архей, геномы дрожжей, червя, мухи, мыши и человека. Минимальное число генов, необходимое для создания простейшей живой клетки (облигатного внутриклеточного паразита), составляет около 470. Минимальное число, необходимое для создания свободноживущей клетки, — около 1700. Типичная грамотрицательная бактерия содержит примерно1500 генов. Штаммы *E. coli* по этому параметру варьируют от 4300 до 5400 генов. Бактериальный ген, в среднем, состоит из 1000 пн и отделен от следующего гена участком около 100 пн. Дрожжи *S. pombe* и *S. serevisiae* содержат 5000 и 6000 генов соответственно.

Несмотря на то что плодовая мушка *D.melanogaster*, на первый взгляд, более сложный организм, а ее геном превышает по размеру геном червя *C.elegans*, у мухи меньше генов (13 600), чем у червя (18 500). Геном растения *Arabidopsis* содержит 25 000 генов, и отсутствие четкой взаимосвязи между размером генома и числом генов для растений можно продемонстрировать тем, что геном риса, который в 4 раза больше, по числу генов превышает геном *Arabidopsis* лишь на 50% (порядка 40 000 генов). Геномы мыши и человека содержат от 20 000 до 30 000 генов, что значительно меньше ранее предполагавшихся значений. Судя по всему, сложность устройства организма может в неменьшей степени зависеть от природы взаимодействия между генами, чем от самого числа генов.

Около 8000 генов являются общими для прокариот и эукариот, и они, скорее всего, вовлечены в выполнение ключевых внутриклеточных функций. Другие 12 000 генов характерны для многоклеточных организмов. Еще 8000 генов встречаются только у животных, и 8000 (в основном отвечающих за функционирование иммунной и нервной систем) — специфичны для позвоночных. У каждого организма (последовательность генома которого определена) только для половины генов установлена определенная функция. Анализ летальных мутаций позволяет сделать вывод, что лишь небольшая часть генов у каждого организма является строго необходимой.

Последовательности эукариотических геномов делят на три группы: уникальные (неповторяющиеся) последовательности, умеренно повторяющиеся последовательности (повторяющиеся небольшое количество раз и разбросанные по геному в виде похожих, но не идентичных последовательностей), и короткие высокоповторяющиеся последовательности (обычно представленные тандемными повторами). Соотношение последовательностей каждого типа является характерной чертой каждого генома. При этом в более крупных геномах доля уникальных последовательностей снижается. Почти половина генома человека состоит из высокоповторяющихся последовательностей, большинство из которых соответствует последовательностям транспозонов. Гены, кодирующие белки, большей частью представляют собой уникальные последовательности ДНК. Дина уникальных регионов ДНК значительно лучше отражает сложность организма, нежели длина генома в целом. Максимальная длина неповторяющейся ДНК (для современных живых систем) достигает 2×10^9 пн.

Уровень экспресси генов варьирует в широком диапазоне значений. В клетке может содержаться 10^5 копий мРНК единственного превалирующего гена (например, кодирующего белок, являющийся главным продуктом данной клетки), 10^3 копий каждой из нескольких (порядка 10) среднепредставленных мРНК, и менее 10 копий каждой мРНК для >10 000 слабоэкспрессирующихся генов. Перекрывание между популяциями мРНК для разных типов клеток может быть весьма значительным; подавляющее большинство мРНК присутствует во всех типах клеток.

Неменделевское наследование объясняется наличием генома в органеллах. Митохондрии и хлоропласты представляют собой двумембранные органеллы, в которых синтезируется лишь часть составляющих их белков, в то время как другая часть импортируется из цитоплазмы. Геномы органелл обычно представлены кольцевой молекулой ДНК, кодирующей все РНК, необходимые для собственной белок-синтезирующей системы, и некоторую часть белков.

Митохондриальные геномы варьируют от 16 тпн (в митохондриях млекопитающих) до 570 тпн (в митохондриях высших растений). Считается, что более крупные митохондриальные геномы кодируют дополнительные функции. Геномы хлоропластов варьируют по размеру от 120 до 200 тпн. Те из них, чья последовательность определена, имеют схожую организацию и кодируют примерно одинаковый набор функций. И в митохондриях, и в хлоропластах многие белковые комплексы одновременно включают субъединицы, синтезированные в пределах органеллы, и субъединицы, импортированные из цитоплазмы.

Митохондриальная ДНК (мтДНК) млекопитающих транскрибируется в единственный протяженный РНК-предшественник, а отдельные мРНК формируются за счет процессинга этой пре-РНК. В мтДНК дрожжей довольно часто встречаются перестройки. Кроме того, между митохондриальными геномами (равно как и между геномами хлоропластов) обнаружена рекомбинация. Существует перенос регионов геномной ДНК хлоропластов и митохондрий в ядерный геном.

Литература

5.2 Число генов у разных бактерий варьирует в десятки раз

Обзор

Bentley, S. D. and Parkhill, J. (2004). Comparative genomic structure of prokaryotes. *Annu. Rev. Genet.* 38, 771–792.

Hacker, J. and Kaper, J. B. (2000). Pathogenicity islands and the evolution of microbes. *Annu. Rev. Microbiol.* 54, 641–679.

Исследование

Blattner, F. R. et al. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* 277, 1453–1474.

Decke[illegible] al. (1998). The complete genome of the [illegible]philic bacterium *Aquifex aeolicus. Nature* 392, [illegible].

Galibert, F. et al. (2001). The composite genome of the legume symbiont *Sinorhizobium meliloti. Science* 293, 668–672.

5.3 Для нескольких эукариотических организмов известно общее число генов

Исследование

Adams, M. D. et al. (2000). The genome sequence of *D. melanogaster. Science* 287, 2185–2195.

Arabidopsis Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana. Nature* 408, 796–815.

C. elegans Sequencing Consortium (1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 282, 2012–2022.

Duffy, A., and Grof, P. (2001). Psychiatric diagnoses in the context of genetic studies of bipolar disorder. *Bipolar Disord.* 3, 270–275.

Dujon, B. et al. (1994). Complete DNA sequence of yeast chromosome XI. *Nature* 369, 371–378.

Goff, S. A. et al. (2002). A draft sequence of the rice genome (*Oryza sativa L. ssp. japonica*). *Science* 296, 92–114.

Johnston, M. et al. (1994). Complete nucleotide sequence of *S. cerevisiae* chromosome VIII. *Science* 265, 2077–2082.

Kellis, M., Patterson, N., Endrizzi, M., Birren, B., and Lander, E. S. (2003). Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* 423, 241–254.

Oliver, S. G. et al. (1992). The complete DNA sequence of yeast chromosome III. *Nature* 357, 38–46.

Wilson, R. et al. (1994). 22 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans. Nature* 368, 32–38.

Wood, V. et al. (2002). The genome sequence of *S. pombe. Nature* 415, 871–880.

5.4 Сколько существует различных типов генов?

Ссылка

Rual, J. F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G. F., Gibbons, F. D., Dreze, M., Ayivi-Guedehoussou, N., et al. (2005). Towards a proteome-scale map of the human protein–protein interaction network. *Nature* 437, 1173–1178.

Обзор

Aebersold, R. and Mann, M. (2003). Mass spectrometry-based proteomics. *Nature* 422, 198–207.

Hanash, S. (2003). Disease proteomics. *Nature* 422, 226–232.

Phizicky, E., Bastiaens, P. I., Zhu, H., Snyder, M., and Fields, S. (2003). Protein analysis on a proteomic scale. *Nature* 422, 208–215.

Sali, A., Glaeser, R., Earnest, T., and Baumeister, W. (2003). From words to literature in structural proteomics. *Nature* 422, 216–225.

Исследование

Agarwal, S., Heyman, J. A., Matson, S., Heidtman, M., Piccirillo, S., Umansky, L., Drawid, A., Jansen, R., Liu, Y., Miller, P., Gerstein, M., Roeder, G. S., and Snyder, M. (2002). Subcellular localization of the yeast proteome. *Genes Dev.* 16, 707–719.

Arabidopsis Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana. Nature* 408, 796–815.

Gavin, A. C. et al. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415, 141–147.

Ho, Y. et al. (2002). Systematic identification of protein complexes in *S. cerevisiae* by mass spectrometry. *Nature* 415, 180–183.

Rubin, G. M. et al. (2000). Comparative genomics of the eukaryotes. *Science* 287, 2204–2215.

Uetz, P. et al. (2000). A comprehensive analysis of protein–protein interactions in *S. cerevisiae. Nature* 403, 623–630.

Venter, J. C. et al. (2001). The sequence of the human genome. *Science* 291, 1304–1350.

5.5 В геноме человека меньше генов, чем кажется

Исследование

Clark, A. G. et al. (2003). Inferring nonneutral evolution from human–chimp–mouse orthologous gene trios. *Science* 302, 1960–1963.

Hogenesch, J. B., Ching, K. A., Batalov, S., Su, A. I., Walker, J. R., Zhou, Y., Kay, S. A., Schultz, P. G., and Cooke, M. P. (2001). A comparison of the Celera and Ensembl predicted gene sets reveals little overlap in novel genes. *Cell* 106, 413–415.

International Human Genome Sequencing Consortium (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860–921.

International Human Genome Sequencing Consortium (2004). Finishing the euchromatic sequence of the human genome. *Nature* 431, 931–945.

Venter, J. C. et al. (2001). The sequence of the human genome. *Science* 291, 1304–1350.

Waterston et al. (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature* 420, 520–562.

5.6 Как гены и прочие последовательности распределены по геному?

Ссылка

Nusbaum, C., Cody, M. C., Borowsky, M. L., Kamal, M., Kodira, C. D., Taylor, T. D., Whittaker, C. A., Chang, J. L., Cuomo, C. A., Dewar, K., et al. (2005). DNA sequence and analysis of human chromosome 18. *Nature* 437, 551–555.

5.7 В Y-хромосоме есть несколько генов, специфичных для самцов

Исследование

Skaletsky, H. et al. (2003). The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* 423, 825–837.

5.8 Усложнение организмов в ходе эволюции происходило за счет появления у генов новых функций

Ссылка

The Chimpanzee Sequencing and Analysis Consortium (2005). Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* 437, 69–87.

5.9 Какая часть генов действительно необходима?

Исследование

Giaever et al. (2002). Functional profiling of the *S. cerevisiae* genome. *Nature* 418, 387–391.

Goebl, M. G. and Petes, T. D. (1986). Most of the yeast genomic sequences are not essential for cell growth and division. *Cell* 46, 983–992.

Hutchison, C. A. et al. (1999). Global transposon mutagenesis and a minimal mycoplasma genome. *Science* 286, 2165–2169.

Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., Welchman, D. P., Zipperlen, P., and Ahringer, J. (2003). Systematic functional analysis of the *C. elegans* genome using RNAi. *Nature* 421, 231–237.

Tong, A. H. et al. (2004). Global mapping of the yeast genetic interaction network. *Science* 303, 808–813.

5.10 Гены отличаются по степени экспрессии

Исследование

Hastie, N. B. and Bishop, J. O. (1976). The expression of three abundance classes of mRNA in mouse tissues. *Cell* 9, 761–774.

5.12 Число экспрессирующихся генов можно определить за один подход

Обзор

Mikos, G. L. G. and Rubin, G. M. (1996). The role of the genome project in determining gene function: insights from model organisms. *Cell* 86, 521–529.

Young, R. A. (2000). Biomedical discovery with DNA arrays. *Cell* 102, 9–15.

Исследование

Holstege, F. C. P. et al. (1998). Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* 95, 717–728.

Hughes, T. R., Marton, M. J., Jones, A. R., Roberts, C. J., Stoughton, R., Armour, C. D., Bennett, H. A., Coffey, E., Dai, H., He, Y. D., Kidd, M. J., King, A. M., Meyer, M. R., Slade, D., Lum, P. Y., Stepaniants, S. B., Shoemaker, D. D. et al. (2000). Functional discovery via a compendium of expression profiles. *Cell* 102, 109–126.

Stolc, V. et al. (2004). A gene expression map for the euchromatic genome of *Drosophila melanogaster. Science* 306, 655–660.

Velculescu, V. E. et al. (1997). Characterization of the yeast transcriptosome. *Cell* 88, 243–251.

Кластеры и повторы

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

6.1 Введение

6.2 Дупликация генов — основная движущая сила эволюции

- С накоплением различий дуплицированные локусы могут превращаться либо в новые гены, либо в неактивные копии.

6.3 Кластеры генов глобинов возникли путем дупликации и дивергенции

- Все глобиновые гены возникли путем дупликации и накопления мутаций. Все они происходят от общего предка, содержащего три экзона.
- Предковый ген дал начало миоглобину, леггемоглобину, а также α- и β-глобинам.
- Разделение α- и β-глобиновых генов произошло на ранней стадии эволюции позвоночных, после чего дупликации привели к появлению обширных кластеров α- и β-подобных генов.
- С момента возникновения инактивирующей ген мутации дальнейший процесс накопления мутаций может привести к формированию псевдогена, гомологичного активному гену (генам).

6.4 Дивергенция последовательностей является основой эволюционных часов

- Последовательности гомологичных генов у разных видов различаются сайтами замен (в которых мутация привела к замене аминокислоты) и молчащими сайтами (мутации в которых не приводят к изменению аминокислотной последовательности белка).
- Мутации в молчащих сайтах накапливаются примерно в 10 раз быстрее, чем в сайтах замен.
- Эволюционные различия двух белков выражают через процент отличающихся аминокислот.
- После разделения генов мутации накапливаются в них с более или менее равной скоростью, таким образом, дивергенция между любой парой глобиновых последовательностей пропорциональна времени, прошедшему с момента их разделения.

6.5 Дивергенция повторяющихся последовательностей отражает скорость накопления нейтральных замен

- В геноме мышей среднегодовая скорость возникновения нейтральных замен выше, чем в геноме человека.

6.6 Псевдогены — тупики эволюции

- Псевдогены не обладают кодирующей функцией, однако анализ последовательностей позволяет определить их родство с существующими функциональными генами. Псевдогены возникают путем аккумуляции мутаций в (некогда) функциональных генах.

6.7 Неравный кроссинговер приводит к реорганизации генетических кластеров

- Если геном содержит кластер генов со схожими последовательностями, спаривание неаллельных генов может приводить к неравному кроссинговеру. Это ведет к формированию делеции в одной из рекомбинантных хромосом и соответствующей дупликации в другой.
- Различные талассемии вызваны делециями, элиминирующими α- или β-глобиновые гены. Тяжесть заболевания у данного индивида зависит от места делеции.

6.8 Гены рРНК формируют тандемные повторы

- Рибосомальная РНК кодируется большим числом идентичных тандемно повторяющихся генов, формирующих один или несколько кластеров.
- Каждый кластер рДНК организован таким образом, что кодирующие предшественник рРНК транскрипционные единицы перемежаются нетранскрибируемыми спейсерами.

6.9 Множественные гены рРНК сохраняют постоянство своей последовательности

- Все гены, расположенные в кластере рДНК идентичны по своей последовательности.
- Каждый нетранскрибируемый спейсер состоит из более коротких повторяющихся единиц, число которых варьирует таким образом, что длины отдельных спейсеров различаются.

6.10 Рекомбинационная фиксация может поддерживать идентичность повторов

- Неравный кроссинговер изменяет размер кластеров тандемных повторов.
- Отдельные повторяющиеся единицы могут быть элиминированы из кластера или вынесены за его пределы.

6.11 Сателлитная ДНК часто расположена в гетерохроматине

- Высокоповторяющаяся ДНК состоит из очень коротких повторяющихся последовательностей и не несет кодирующей функции.
- Она встречается в виде крупных блоков, которые могут иметь различные физико-химические свойства.
- Она часто является основной составной частью центромерного гетерохроматина.

6.12 Сателлитная ДНК членистоногих состоит из очень коротких идентичных повторов

- Повторяющиеся единицы сателлитной ДНК членистоногих составляют всего несколько нуклеотидов в длину. Большинство копий последовательности идентичны.

6.13 Сателлитная ДНК млекопитающих состоит из иерархических повторов

- Сателлитная ДНК мышей эволюционировала путем дупликации и мутирования исходной короткой повторяющейся единицы в основную повторяющуюся единицу размером 234 пн, в половине, четверти, и восьмой части которой можно распознать исходный повтор.

6.14 Мини-сателлиты можно использовать при генетическом картировании

- Вариабельность длины микросателлитов и мини-сателлитов отдельных геномов можно использовать для определения степени их биологического родства. Для этого достаточно показать, что 50% исследованных микросателлитов и мини-сателлитов индивидуума получены от определенного родителя.

6.15 Резюме

6.1 Введение

Набор генов, произошедших путем многократных дупликаций общего предшественника с последующим накоплением отличий, называют **семейством генов**. Члены семейства могут быть собраны в одном месте генома или распределены по разным хромосомам. Исследования показали, что многие гены представлены семействами. 25 000 генов человека распределены примерно по 15 000 семейств, таким образом, в среднем, каждый ген представлен в геноме двумя «родственниками» (рис. 5.7). Семейства сильно отличаются по степени родства между составляющими их членами. Некоторые состоят из множества идентичных представителей, а некоторые включают членов, связанных очень дальним родством. Обычно гены схожи последовательностями экзонов и отличаются интронами (см. разд. 3.5 «Последовательности экзонов консервативны, а интронов — вариабельны»). Гены могут быть гомологичны лишь частью своих экзонов, в остальном оставаясь уникальными (см. разд. 3.9 «Некоторые экзоны напоминают последовательности, кодирующие целые белки»).

Событием, приводящим к возникновению родственных экзонов или целых генов, служит дупликация, формирующая в пределах генома копию некой последовательности. Тандемные дупликации (в которых пары копий расположены рядом) могут возникать в результате ошибок репликации или рекомбинации. Разделение этих пар может осуществляться путем **транслокации**, когда материал из одной хромосомы переносится в другую. Копия может возникнуть и на новом месте в геноме. Это может происходить в результате транспозиции, связанной с копированием участка ДНК, расположенного в непосредственной близости от транспозона. Дуплицироваться могут как целые гены, так и несколько экзонов или даже единичные экзоны. При дупликации целого гена, формируются две копии, выполняющие одинаковую функцию, однако впоследствии за счет возникновения мутаций эти копии накапливают различия.

Тесно связанные по своей структуре члены семейства обычно выполняют сходные (или даже идентичные) функции, хотя они могут экспрессироваться в различное время или в разных типах клеток. Так, в эмбриональных и взрослых эритроцитах экспрессируются различные белки семейства глобинов, тогда как различные актины помимо клеток мускулатуры используются в широком спектре других клеток. Если гены сильно различаются или если они гомологичны лишь по некоторым экзонам, кодируемые ими белки могут иметь разные функции.

Некоторые семейства генов состоят из идентичных представителей. Объединение в кластеры является предпосылкой для сохранения тождественности генов, хотя гены, принадлежащие одному кластеру, не обязательно идентичны. Размеры возникших за счет дупликаций **генных кластеров** могут варьировать в диапазоне от двух смежных генов до сотен идентичных генов, тандемно расположенных на участке ДНК. Многократные тандемные повторы гена могут возникать в тех случаях, когда кодируемый ими продукт нужен в большом количестве. Типичным примером служат гены рРНК или белков гистонов. Такие ситуации возникают благодаря сохранению идентичности генов и давлению естественного отбора.

Кластеры генов дают нам возможность оценить влияние сил, вовлеченных в эволюцию генома, на более крупных участках, нежели отдельные гены. Дуплицированные последовательности (в особенности расположенные рядом друг с другом) являются основой для дальнейшей эволюции путем рекомбинации. Популяция эволюционирует за счет классической рекомбинации с точным кроссинговером (см. РИС. 6.1 и 6.2). Рекомбинантные хромосомы имеют ту же организацию, что и родительские. Они содержат те же локусы, расположенные в том же порядке, однако аллели в них представлены в других комбинациях. Дальше в работу с этим «сырым» материалом включается естественный отбор. Дупликации являются источником случайных отклонений от правильного течения рекомбинационных процессов, приводя к изменению не только комбинации аллелей, но и смысла генов.

Неравный кроссинговер (известный также как **нереципрокная рекомбинация**) представляет собой рекомбинацию, происходящую между двумя него-



РИС. 6.1 Образование хиазмы предшествует появлению рекомбинантов

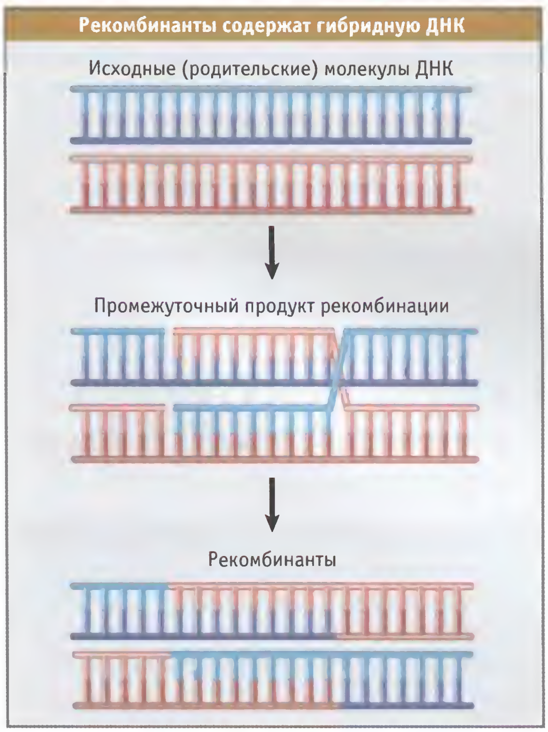


РИС. 6.2 При рекомбинации происходит взаимодействие комплементарных цепей из разных исходных дуплексов ДНК

мологичными сайтами. Возможность такого события обусловлена наличием повторяющихся последовательностей. На РИС. 6.3 показано, что при неравном кроссинговере одна копия повтора (расположенная на одной из гомологичных хромосом) вместо того, чтобы рекомбинировать с гомологичным регионом второй хромосомы, может рекомбинировать по месту расположения другой копии повтора. При рекомбинации это увеличивает число повторов на одной хромосоме и уменьшает на другой. На самом деле одна из рекомбинантных хромосом получает делецию, а другая — инсерцию. Этот механизм ответственен за эволюцию кластеров повторяющихся последовательностей. Мы можем отследить этот процесс по увеличению или уменьшению длины повторяющихся участков как на примере генных кластеров, так и для регионов высокоповторяющейся ДНК.

Фракция высокоповторяющихся последовательностей генома представлена множеством тандемно расположенных копий коротких повторяющихся единиц. Эта фракция ДНК обладает необычными свойствами. Во-первых, она выявляется в виде отдельного пика при центрифугировании ДНК в градиенте плотности (поэтому данная фракция получила название **сателлитной ДНК**). Кроме того, высокоповторяющиеся последовательности часто ассоциированы с неактивными регионами хромосомы, в частности с центромерами (в которых расположены места присоединения к митотическому или мейотическому веретену деления). Из-за повторяющейся структуры сателлитная ДНК демонстрирует сходное (в эволюционном смысле) поведение с кластерами тандемных генов. Кроме собственно сателлитных последовательностей, существуют более короткие повторы в ДНК, называемые **мини-сателлитами** (и в целом ведущие себя так же, как сателлитная ДНК). Мини-сателлиты хорошо подходят для демонстрации различий между отдельными геномами, что позволяет успешно использовать их при картировании.

Все описанные события изменяют строение генома и довольно редки, однако они играют важную роль в процессе эволюции.

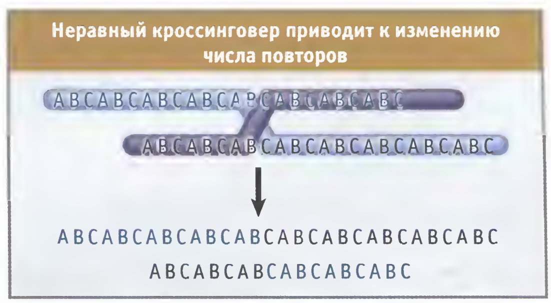


РИС. 6.3 Неравный кроссинговер является следствием спаривания между неэквивалентными участками в случаях, когда часть хромосомы, где образуется хиазма, состоит из повторов ДНК. На рисунке, где единичный повтор обозначен как ABC, видно, что третий повтор на голубой хромосоме по ошибке пришел в соответствие с первым повтором на черной хромосоме. (Поскольку область взаимодействия хромосом достаточно обширна, соседние повторы тоже пришли в соответствие «не с теми» повторами на другой хромосоме.) Произошел сдвиг хромосом друг относительно друга. После завершения кроссинговера получатся хромосомы, несущие шесть и десять повторов (вместо двух по восемь, как при нормальном кроссинговере)

6.2 Дупликация генов — основная движущая сила эволюции

Ключевые понятия:

- С накоплением различий дуплицированные локусы могут превращаться либо в новые гены, либо в неактивные копии.

Экзоны — это модули для построения генов, испытанные в процессе эволюции в различных комбинациях. С одной стороны, любой отдельно взятый экзон некоторого гена может быть дуплицирован и использован при «построении» нового гена. С другой стороны, дупликация может затрагивать весь ген (включая все его экзоны и интроны). В этом случае в одной из копий, вследствие отсутствия негативного влияния естественного отбора, могут накапливаться мутации. Впоследствии эта копия может измениться настолько, что приобретет новую функцию (например, новую активность или способность экспрессироваться в другое время или в другом месте).

На **РИС. 6.4** приведены современные представления о динамике этих процессов. Вероятность дупликации некоторого отдельно взятого гена за миллион лет равна примерно 1%. Мутации, возникающие в каждой из копий после дупликации, формируют различия между ними. Мутации накапливаются со скоростью ~ 0,1% за миллион лет (см. разд. 6.4 «Дивергенция последовательностей является основой эволюционных часов»).

Организму нет необходимости сохранять две идентичные копии гена. Накапливающие различия копии могут развиваться двумя путями:

- Оба гена становятся важными для организма. Это может произойти в том случае, если накопленные отличия приводят к синтезу белка с новыми свойствами или если новые гены приобретают способность экспрессироваться в разное время или в различных местах.
- Если этого не происходит, один из генов, скорее всего, будет удален из генома (такой делеционный мутант не будет испытывать давления негативного отбора). Обычно для этого необходимо около 4 млн лет. Выбор копии, которая станет неактивной, носит случайный характер (и это может быть основой формирования различий между отдельными особями, а в случае потери активности различных копий в разных популяциях, даже приводить к образованию новых видов).

Геном человека примерно на 5% состоит из дупликаций сегментов, варьирующих по длине от 10 до 300 тпн. Эти дупликации возникли сравнительно недавно (т. е. прошедшего времени оказалось недостаточно для того, чтобы окончательно стереть черты, указывающие на их родство). Они содержат стандартную для генома долю экспрессируемых экзонов (около 6%), что указывает на независимое от сущности генетического материала возникновение дупликаций. Гены, лежащие в этих дупликациях, представляют особый интерес, поскольку они эволюционировали сравнительно недавно и, таким образом, могли играть важную роль в событиях, происходивших в новейшей эволюции (например, при происхождении человека от обезьяны).

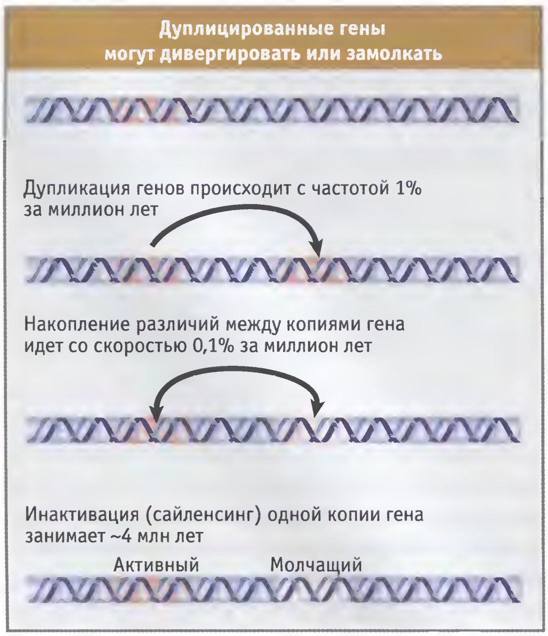


РИС. 6.4 После того как ген оказался дуплицирован, между двумя его копиями могут возникать и накапливаться различия. Копии гена могут приобрести разные функции, одна из них может быть инактивирована

6.3 Кластеры генов глобинов возникли путем дупликации и дивергенции

Ключевые понятия:

- Все глобиновые гены возникли путем дупликации и накопления мутаций. Все они происходят от общего предка, содержащего три экзона.
- Предковый ген дал начало миоглобину, леггемоглобину, а также α- и β-глобинам.
- Разделение α- и β-глобиновых генов произошло на ранней стадии эволюции позвоночных, после чего дупликации привели к появлению обширных кластеров α- и β-подобных генов.
- С момента возникновения инактивирующей ген мутации, дальнейший процесс накопления мутаций может привести к формированию псевдогена, гомологичного активному гену (генам).

Чаще всего дупликация приводит к формированию второй копии гена вблизи от первой. В некоторых случаях обе копии так и остаются рядом, и тогда следующая дупликация может сформировать уже целый кластер родственных генов. Самым хорошо охарактеризованным примером кластера генов являются гены глобинов, составляю-

щие древнее семейство, связанное с выполнением одной из самых важных для представителей царства животных функцией: транспортом кислорода с кровотоком.

Глобиновый тетрамер является главной функциональной частью эритроцита. Вместе с гемом (группой, связывающей железо) он формирует гемоглобин. У представителей разных видов работающие гены глобинов схожи по своему строению (как было показано ранее на рис. 3.7, они включают три экзона). Основываясь на данном сходстве, можно сделать вывод, что все они произошли от общего предка. Изучая развитие отдельных глобиновых генов на внутри- и межвидовом уровнях, можно получить представление о механизмах, вовлеченных в эволюцию семейств генов.

В эритроцитах взрослых организмов глобиновый тетрамер состоит из двух идентичных α-цепей и двух идентичных β-цепей. Однако в клетках крови эмбриона тетрамеры гемоглобина отличаются от формы, работающей в клетках взрослого организма. У эмбриона каждый тетрамер содержит две идентичные α-подобные цепи и две идентичные β-подобные цепи, каждая из которых гомологична соответствующему «взрослому» полипептиду, которым впоследствии и заменяется. Это пример контролируемого развития, при котором происходит успешное включение и выключение генов, приводящее к синтезу альтернативных продуктов, выполняющих одну и ту же функцию, но в различное время.

Деление цепей глобинов на α-подобные и β-подобные отражает особенности строения генов. Каждый тип глобина закодирован генами, принадлежащими одному кластеру. Организация двух глобиновых кластеров в геноме высших приматов изображена на **РИС. 6.5**.

β-кластер, занимающий 50 тпн, содержит пять активных генов (ε, два γ, δ и β) и один неактивный (ψβ). Два гена γ оличаются в своей кодирующей части лишь одной аминокислотой: у варианта G в позиции 136 расположен глицин, а у варианта A — аланин.

Более компактный кластер α занимает 28 тпн и содержит один активный ген ζ, один неактивный ген ζ, два активных и два неактивных гена α, и один ген θ, функция которого неясна. Оба активных гена α кодируют один и тот же белок. Два (или более) идентичных гена, расположенных на одной хромосоме, называются **неаллельными** копиями.

Детали перехода между эмбриональными и взрослыми гемоглобинами варьируют у разных организмов. У человека эритроциты проходят через три стадии: эмбриональную, фетальную, и зрелую. Выделение эмбриональной и зрелой стадий характерно для всех млекопитающих, но число промежуточных стадий варьирует. У человека α-подобные цепи называют альфа и зета, а β-подобные — бета, гамма, дельта и эпсилон. На **РИС. 6.6** показано, как на разных стадиях развития экспрессируются различные цепи.

Первым у человека экспрессируется ген ζ α-подобной цепи, однако вскоре начинает экспрессироваться ген α. В β-пути первыми экспрессируются гены ε и γ, а



РИС. 6.5 Каждое семейство α- и β-подобных глобиновых генов организовано в виде одиночного кластера, в котором есть и функциональные гены, и псевдогены (ψ)

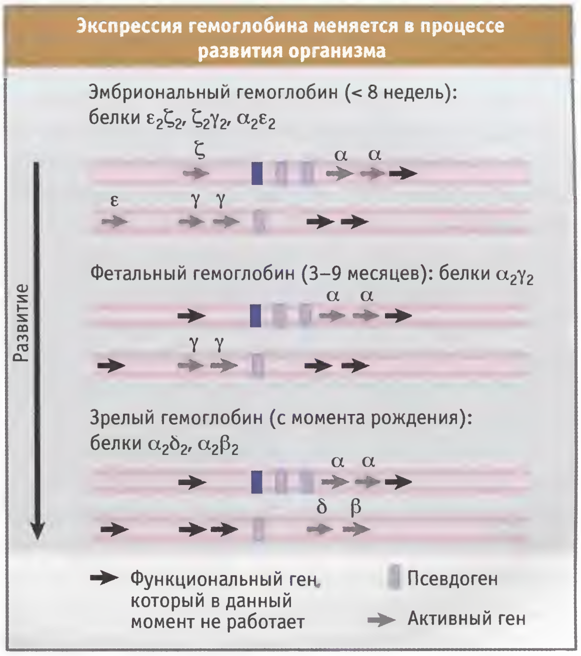


РИС. 6.6 В разные периоды индивидуального развития у человека экспрессируются разные гемоглобиновые гены

δ и β заменяют их позже. У взрослых людей гемоглобин на 97% представлен формой $\alpha_2\beta_2$, на 2% формой $\alpha_2\delta_2$ и на 1% пренатальной формой $\alpha_2\gamma_2$.

В чем смысл отличия эмбриональных и зрелых форм глобинов? Дело в том, что эмбриональные и фетальные формы имеют большее сродство к кислороду. Это необходимо для получения кислорода из крови матери. Понятно, почему подобный эквивалент отсутствует, скажем, у цыплят: у них эмбриональные стадии проходят вне тела матери (в яйце).

Функциональные гены отличаются способностью продуцировать мРНК и, в конечном счете, способностью производить кодируемые ими белки. Молчащие гены не продуцируют белки. Существует несколько причин, по



РИС. 6.7 Кластеры β-глобиновых генов и псевдогенов есть у всех позвоночных. У кролика и курицы в этих кластерах по четыре гена. У мыши он насчитывает семь генов: два гена, экспрессирующихся в раннем эмбриогенезе, один — в позднем эмбриогенезе, два гена зрелой стадии, а также два псевдогена

которым гены становятся неактивными. Чаще всего это происходит из-за нарушения процессов транскрипции и/или трансляции. Такие гены называют **псевдогенами** и обозначают буквой ψ. Подобная организация глобиновых кластеров встречается у многих позвоночных, но, как показано на РИС. 6.7, типы, число и порядок генов у них могут варьировать. Любой кластер содержит и эмбриональные, и характерные для взрослых форм гены. Суммарная длина глобиновых кластеров варьирует в широких пределах. Наибольшую (среди изученных) длину кластеров имеют козы, у которых основной кластер, состоящий из четырех генов, дуплицирован дважды. Причем в каждом из «дубликатов» активные гены и псевдогены распределены по-разному, что иллюстрирует вероятностную природу превращения одной из копий дуплицированного гена в неактивную форму.

Приведенное описание глобиновых кластеров позволяет нам сделать важный вывод: *в семействе генов может быть больше активных и неактивных представителей, чем можно предположить на основе анализа белков*. Дополнительные активные гены могут представлять собой копии, кодирующие идентичные полипептиды, или гомологичные (но при этом отличающиеся) известным белки (которые обычно нарабатываются кратковременно или в небольших количествах).

Возвращаясь к вопросу о том, сколько ДНК необходимо для того, чтобы закодировать определенную функцию, мы видим, что для кодирования β-подобных глобинов различным млекопитающим нужно от 20 до 120 тпн. Это значительно больше, чем можно было бы предположить, основываясь только лишь на подробном анализе известных белков β-глобинов или даже принимая во внимание отдельные гены. Заметим, что кластеры такого типа встречаются нечасто; большинство генов существует в виде уникальных локусов.

На примере организации генов глобинов у представителей различных видов можно проследить возникновение современных кластеров генов глобинов от общего предшественника. Нынешние представления об эволюционном происхождении глобинов изображены на РИС. 6.8.

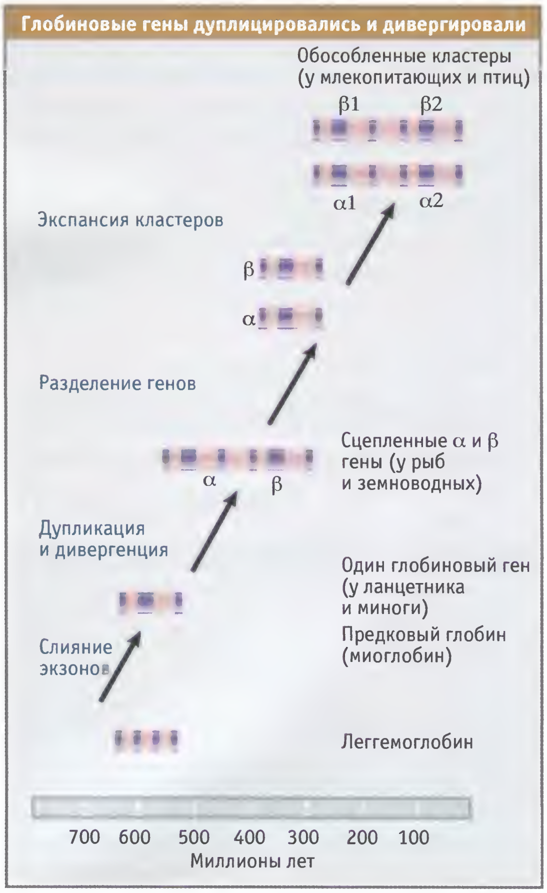


РИС. 6.8 Все глобиновые гены происходят от одного предкового гена (претерпевшего серию дупликаций, транспозиций и мутаций)

За предковую форму можно принять родственный глобинам ген леггемоглобина растений. Среди ныне существующих глобинов наиболее близкой к предковой форме является последовательность единственной цепи гена миоглобина млекопитающих, отделившаяся от линии глобинов около 800 миллионов лет назад. Организация гена миоглобина аналогична таковой у других глобиновых генов (трехэкзонная структура может служить дополнительным доказательством общности происхождения).

У некоторых «примитивных рыб» (миксин и миног) присутствует лишь один тип глобиновых цепей, т. е. в ходе эволюции они должны были отделиться до того, как предок глобиновых генов был дуплицирован и дал начало вариантам α и β. Вероятнее всего, это произошло около 500 миллионов лет назад во время эволюции костистых рыб.

Следующую стадию эволюции генов глобинов можно обнаружить у лягушки *Xenopus laevis*, имеющей два глобиновых кластера. Однако каждый из кластеров у нее содержит как α-, так и β-гены, причем и личиночного, и взрослого типов. Таким образом, кластер должен был эволюционировать путем дупликации сцепленных α–β-пар с последующим разделением отдельных копий, и только после этого весь кластер был дуплицирован.

Земноводные отделились от линии птиц и млекопитающих около 350 миллионов лет назад. Следовательно, разделение на α- и β-глобиновые гены стало результатом геномной перестройки у некоего предшественника птиц и млекопитающих уже после этого события. Возможно, это произошло во времена ранней эволюции позвоночных. И у птиц, и у млекопитающих для α- и β-глобинов существуют отдельные кластеры, следовательно, физическое разделение α- и β-генов должно было произойти до того, как млекопитающие и птицы разошлись от общего предка, то есть около 270 миллионов лет назад.

Как мы увидим из приведенного в следующем разделе описания дивергенции отдельных генов, изменения внутри α- и β-кластеров возникли позже.

6.4 Дивергенция последовательностей является основой эволюционных часов

Ключевые понятия:

- Последовательности гомологичных генов у разных видов различаются сайтами замен (в которых мутация привела к замене аминокислоты) и молчащими сайтами (мутации в которых не приводят к изменению аминокислотной последовательности белка).
- Мутации в молчащих сайтах накапливаются примерно в 10 раз быстрее, чем в сайтах замен.
- Эволюционные различия двух белков выражают через процент отличающихся аминокислот.
- После разделения генов мутации накапливаются в них с более или менее равной скоростью, таким образом, дивергенция между любой парой глобиновых последовательностей пропорциональна времени, прошедшему с момента их разделения.

Главным механизмом изменения последовательности белка является медленное накопление незначительных мутаций. Точечные мутации, а также небольшие вставки и делеции возникают случайно с более или менее равной частотой во всех регионах генома. Исключением из этого правила являются так называемые «горячие точки», в которых мутации возникают чаще обычного. Большая часть изменяющих аминокислотную последовательность мутаций относится к классу делеций, которые элиминируются из популяции под действием естественного отбора.

Однако некоторые мутации обладают положительным эффектом. Возникая, они быстро распространяются в популяции, постепенно заменяя исходный аллель. Когда новый вариант замещает предыдущую версию гена, говорят, что он *зафиксировался* в популяции.

Спорным остается вопрос о частоте так называемых **нейтральных** мутаций, приводящих к аминокислотной замене, не влияющей на функцию белка, но при этом способных накапливаться в результате **случайного дрейфа** и **фиксации**.

Скорость накопления мутационных изменений для каждого белка своя и по крайней мере отчасти она зависит от способности белка сохранять свою функцию при замене отдельных аминокислот. Внутривидовая эволюция белков идет путем возникновения мутантов с последующим элиминированием или фиксацией в пределах совокупности организмов, способных скрещиваться друг с другом. Важно помнить, что при анализе совокупности аллелей любого гена, мы видим только те варианты, которым удалось сохраниться в эволюции. Множественные варианты генов либо стабильны (поскольку ни один из них не дает явного селективного преимущества), либо некоторые из аллелей находятся в переходной фазе замещения.

После разделения одного вида на два новых, каждый из вновь образованных видов становится эволюционно независим (с генетической точки зрения). Сравнивая соответствующие белки у представителей двух видов, мы видим отличия, накопившиеся у них *с тех пор, как их предки перестали скрещиваться*. Некоторые белки очень консервативны и практически неизменны у представителей разных видов. Это указывает на то, что практически любое их изменение приводит к потере белком своей функции, и такие варианты элиминируются.

Различия между двумя белками выражают через их **дивергенцию** — процент позиций, по которым различаются их аминокислотные последовательности. Дивергенция между белками может отличаться от дивергенции соответствующих им последовательностей нуклеиновых кислот. Причиной этих отличий является вырожденность генетического кода, благодаря чему изменение нуклеотида не всегда влияет на значение кодона.

Нуклеотидную последовательность кодирующего региона можно разделить на потенциальные **сайты замен** и **молчащие сайты**:

- Мутация, расположенная в сайте замен, приводит к замене аминокислоты. Эффект, производимый этой мутацией (отрицательный, нейтральный или благо-

приятный) зависит от результата, к которому приводит замена соответствующей аминокислоты.

- В молчащих сайтах мутация просто заменяет один синонимичный кодон на другой (то есть в самом белке никаких изменений не происходит). Обычно 75% кодирующей последовательности составляют сайты замен, а остальные 25% — молчащие сайты.

Кроме кодирующей последовательности ген содержит нетранслируемые регионы. Расположенные в них мутации также потенциально нейтральны, независимо от эффектов, которые они оказывают на вторичную структуру белка или на (обычно довольно короткие) регуляторные сигналы.

Несмотря на то что молчащие мутации нейтральны по отношению к белку, они могут влиять на экспрессию гена путем изменения последовательности РНК. Например, изменения вторичной структуры мРНК могут влиять на транскрипцию, процессинг или трансляцию. Также возможна ситуация, когда изменения в синонимичных кодонах приводят к связыванию с различными тРНК (с разной представленностью в цитоплазме), влияя, таким образом, на эффективность трансляции.

Мутации, расположенные в сайтах замен, должны соответствовать аминокислотной дивергенции (определяемой как процент замен в последовательности белка). Дивергенция нуклеиновых кислот в сайтах замен на уровне 0,45% соответствует 1% дивергенции аминокислот (из расчета, что в среднем на кодон приходится 2,25 сайтов замен). На практике при измерении дивергенции отличия, возникавшие в одном и том же кодоне несколько раз в ходе эволюции, не видны (обычно на это делается поправка).

На примере человеческих β- и δ-глобиновых цепей можно увидеть 10 отличий в последовательности из 146 аминокислот, что соответствует уровню дивергенции в 6,9%. Последовательность ДНК длиной в 441 нуклеотид несет 31 отличие. Однако изменения распределены по молчащим сайтам и сайтам замен неравномерно. Из 111 молчащих сайтов отличаются 20, а из 330 сайтов замен — всего 11. С учетом этого обстоятельства можно получить (корректную) дивергенцию на уровне 3,7% для сайтов замен и на уровне 32% для молчащих сайтов — значения, отличающиеся почти на порядок.

Такая сильная разница в значении дивергенции для молчащих сайтов и сайтов замен показывает, что на позиции нуклеотидов, влияющих на строение белка, оказывается большее давление, чем на позиции нуклеотидов, не влияющих на его структуру. Можно предположить, что нейтральными являются лишь немногие из аминокислотных замен в этих белках.

Примем частоту мутаций в молчащих сайтах за величину, отражающую базовую скорость фиксации мутаций (при допущении, что изменения в молчащих сайтах вовсе не испытывают влияния отбора). Тогда с момента разделения β- и δ-генов в 32% из 330 сайтов замен должно было возникнуть 105 мутаций, все кроме 11 из которых были элиминированы, а это означает, что около 90% мутаций «не выжили».

Дивергенция между любой парой последовательностей глобинов (в той или иной степени) пропорциональна времени, прошедшему с момента их разделения. Это позволяет нам ввести понятие **эволюционных часов**, которые измеряют среднюю скорость накопления мутаций в ходе эволюционного развития данного белка.

Скорость дивергенции можно выразить через процент отличий за миллион лет или через обратную величину — единицу эволюционного времени (unit evolutionary period, UEP), показывающую время (в миллионах лет), для дивергенции на 1%. Будучи основаны на попарном межвидовом сравнении (вспомним сложности, возникавшие при установлении реального времени происхождения вида), эти часы можно применять к родственным генам *в пределах* вида. Выяснив дивергенцию гомологичных генов, можно посчитать, сколько времени прошло с момента их дупликации.

Сравнивая последовательности гомологичных генов в разных видах, можно определить скорость дивергенции как в сайтах замен, так и в молчащих сайтах (РИС. 6.9).

Попарное сравнение демонстрирует, что у млекопитающих, отделившихся около 85 миллионов лет назад, дивергенция в сайтах замен как α-, так и β-глобиновых генов составляет 10%. Это соответствует значению скорости дивергенции в сайтах замен на уровне 0,12% за миллион лет.

Если распространить данное сравнение на гены, разделившиеся раньше, уровень останется неизменным. Например, средняя дивергенция по сайтам замен между соответствующими глобиновыми генами млекопитающих и птиц составляет 23%. Относительно разделения, произошедшего около 270 миллионов лет назад, это дает скорость на уровне 0,09% за миллион лет.

Двигаясь еще дальше, мы можем сравнить α- и β-глобиновые гены в пределах видов. Они накапливали

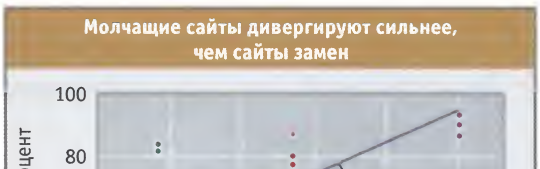


РИС. 6.9 Дивергенция последовательностей ДНК зависит от их расхождения в процессе эволюции. Каждая точка на графике соответствует результату сравнения пары сайтов

отличия, начиная со времени дупликации единственного гена, произошедшей около 500 миллионов лет назад (рис. 6.8). Средняя дивергенция по сайтам замен у них составляет примерно 50%, что дает скорость на уровне 0,1% за миллион лет.

Сводная информация по этим данным представлена на рис. 6.9. На нем показано, что дивергенция по сайтам замен глобиновых генов составляет в среднем около 0,096% за миллион лет (или 10,4 UEP). С учетом ошибки определения времени разделения видов, полученные результаты свидетельствуют в пользу поддержки идеи о существовании линейных часов.

Данные по дивергенции в молчащих сайтах менее однозначны. Видно, что дивергенция в молчащих сайтах сильнее, чем в сайтах замен, и коэффициент, отражающий это отличие, варьирует от 2 до 10. Таким образом, разброс значений дивергенции для молчащих сайтов (оцениваемых при попарном сравнении) слишком велик для того, чтобы использоваться в качестве эволюционных часов (а временные сравнения надо строить на данных по сайтам замен).

Из рис. 6.9 становится понятно, что скорость дивергенции в молчащих сайтах нелинейна относительно времени. *Если предположить, что в момент дупликации дивергенция равна нулю*, можно заметить, что на протяжении первых 100 миллионов лет (с момента дупликации) скорость дивергенции молчащих сайтов будет выше, чем в последующий период. Этот феномен можно объяснить тем, что некая часть (примерно половина) молчащих сайтов быстро (за первые 100 миллионов лет) насыщается мутациями, то есть они ведут себя как истинно нейтральные. Другая же половина молчащих сайтов накапливает мутации гораздо медленнее, со скоростью, приблизительно равной скорости накопления мутаций в сайтах замен. Несмотря на то что эти сайты с точки зрения аминокислотной последовательности являются молчащими, по каким-то другим причинам они испытывают давление отбора.



РИС. 6.10 Дивергенция сайтов замен в парах β-глобиновых генов позволяет восстановить картину эволюции кластера у человека. Схема показывает историю обособления классов глобиновых генов

Подсчет скорости дивергенции можно использовать и для оценки времени разделения генов в пределах вида. У человека отличия в сайтах замен между генами β и δ составляют 3,7%. Значение в 10,4 UEP показывает, что гены должны были разделиться 10,4 × 3,7 = 40 миллионов лет назад — примерно во времена разделения линий, ведущих к обезьянам Нового Света, Старого Света, человекообразным обезьянам и человеку. У всех этих высших приматов есть гены β и δ, из чего можно сделать вывод, что дупликация произошла непосредственно перед разделением этих линий в эволюции.

Пытаясь заглянуть в еще более далекое прошлое, обратимся к дивергенции между сайтами замен генов γ и ε, составляющей 10%, что соотносится с временем разделения этих генов около 100 миллионов лет назад. Таким образом, разделение на эмбриональные и фетальные гены глобинов могло непосредственно предшествовать появлению млекопитающих или происходить одновременно с их формированием.

Строение эволюционного дерева генов глобина человека приведено на РИС. 6.10. Характерные черты, возникшие в результате эволюционных событий, произошедших до появления млекопитающих (например, таких, как отделение β/δ от γ), должны присутствовать у всех млекопитающих. События же, произошедшие после этого (например, разделение β- и δ- генов глобина), должны прослеживаться в отдельных линиях млекопитающих.

У каждого из видов в структуре кластеров можно обнаружить и сравнительно недавние изменения. К ним относятся отличия в числе генов (один зрелый β-глобин у человека и два — у мыши) или в их типе (чаще всего касающиеся наличия отдельных эмбриональных и пренатальных генов).

Таким образом, при наличии достаточного количества данных относительно определенных последовательностей, путем сравнения гомологичных генов у представителей разных видов можно определять связи между таксонами.

6.5 Дивергенция повторяющихся последовательностей отражает скорость возникновения нейтральных замен

Ключевые понятия:

- В геноме мышей среднегодовая скорость возникновения нейтральных замен выше, чем в геноме человека.

Изучение не кодирующих белки последовательностей позволяет наилучшим образом оценить скорость возникновения замен в нейтральных сайтах (термин «ней-

тральный» вместо термина «молчащий» в данном случае выбран для того, чтобы подчеркнуть отсутствие кодирующего потенциала). Так, информативным может быть сравнение общего для человека и мыши семейства повторяющихся элементов генома.

Принцип анализа приведен на РИС. 6.11. Возьмем семейство гомологичных последовательностей, возникших из исходной последовательности путем дупликаций и накопления замен. Предположим, что общую предковую последовательность можно составить из наиболее часто встречающихся в каждой из позиций нуклеотидов. После этого посчитаем дивергенцию для каждого из отдельных членов семейства в виде доли нуклеотидов, отличающихся от последовательности, принятой нами за предковую. В данном случае дивергенция для отдельных представителей варьирует в диапазоне от 0,13 до 0,18 (в среднем — 0,16).

Одно из проанализированных таким способом семейств в геномах человека и мыши происходит от последовательности, потерявшей активность приблизительно во времена разделения ветвей приматов и грызунов (семейство LINES; см. разд. 22.9 «Систематика ретропозонов»). Это означает, что данные последовательности накапливали отличия на протяжении равного для обоих видов промежутка времени, не испытывая при этом давления отбора. У человека средняя дивергенция для членов семейства LINES составляет приблизительно 0,17 замен на сайт, что соответствует скорости в $2{,}2 \times 10^{-9}$ замен на основание за год на протяжении 75 миллионов лет (с момента дупликации). Следует заметить, что в геноме мыши нейтральные замены возникали в два раза быстрее. Дивергенция для них составляет 0,34 замены на сайт, а скорость составляет $4{,}5 \times 10^{-9}$. При этом если посчитать скорость не за год, а за поколение, то у человека она будет выше ($\sim 2{,}2 \times 10^{-8}$ у человека против 10^{-9} у мыши).

Полученные на основе семейства LINES данные, возможно, занижают скорость накопления изменений в геноме мышей. В момент разделения линий человека и мыши скорость изменения в обеих линиях должна была быть равной (а наблюдаемые сегодня отличия накоплены как раз с момента разделения предковых линий). Нынешняя годовая скорость накопления нейтральных замен у мышей, возможно, в 2–3 раза выше, чем среднее значение за период, прошедший с момента разделения предковых линий. Скорость накопления замен отражает баланс между возникновением мутаций и способностью систем репарации исправлять их. Современные отличия в скоростях накопления изменений для человека и мыши можно объяснить различиями в эффективности их репарационных систем.

Сравнение геномов мыши и человека позволяет определить, существуют ли более консервативные области генома или все последовательности в равной мере отличаются друг от друга (пропорционально скорости накопления нейтральных замен). Процент регионов, демонстрирующих признаки консервирующего давления отбора, составляют около 5% генома. Это значительно больше, чем доля последовательностей, кодирующих белки или структурные РНК (всего около 1%). Это означает, что геном содержит даже больше важных некодирующих участков, чем последовательностей, кодирующих белок или РНК. На сегодняшний день изучена лишь небольшая часть таких регуляторных последовательностей. Полученные данные также свидетельствуют в пользу того, что большая часть генома (т. е. оставшиеся примерно 95%) не выполняют каких-либо функций, зависящих от конкретной последовательности нуклеотидов.



РИС. 6.11 Предковую последовательность-консенсус семейства последовательностей можно восстановить. Для этого в каждое положение последовательности ставится нуклеотид, свойственный большинству членов семейства. Дивергенцию можно рассчитать как долю положений, по которым член семейства отличается от консенсуса

6.6 Псевдогены — тупики эволюции

Ключевые понятия:

- Псевдогены не обладают кодирующей функцией, однако анализ последовательностей позволяет определить их родство с существующими функциональными генами. Псевдогены возникают путем аккумуляции мутаций в (некогда) функциональных генах.

Псевдогены (обозначаемые буквой Ψ перед названием гена) характеризуются наличием последовательностей, гомологичных последовательностям функциональных генов, однако при этом они не дают функциональных белков.

Некоторые псевдогены обладают тем же планом строения, что и активные гены. Экзоны и интроны псевдогена расположены в тех же местах, что и у гена-гомолога. Псевдогены возникают путем потери геном активности в результате мутаций, нарушающей одну (или несколько) стадий экспрессии гена. Нарушения могут затрагивать сигналы инициации транскрипции, сплайсинг в местах соединения экзонов и интронов, или провоцировать преждевременную терминацию трансляции.

Обычно псевдоген содержит несколько делеций. Скорее всего, однажды потеряв активность, он начинает беспрепятственно накапливать мутации. Псевдогены, представляющие собой неактивные копии существующих активных генов, были найдены во многих семействах, включая глобины, иммуноглобулины и системы антигенов комплекса гистосовместимости. В большинстве случаев псевдогены расположены вблизи кластера генов-гомологов, часто даже перемежая активные копии генов.

Типичным примером псевдогена является Ψβ2 кролика, наиболее близкий по последовательности активному гену глобина β1 и имеющий с ним общую схему строения. Однако Ψβ2 неактивен. На РИС. 6.12 приведены нарушения, характерные для данного псевдогена. Делеция пары нуклеотидов в 20-м кодоне Ψβ2 приводит к сбою рамки считывания и появлению стоп-кодона чуть дальше делеции. Вдобавок, несколько точечных мутаций изменяют кодоны высококонсервативных для β-глобинов аминокислот. Ни один из двух интронов не содержит распознаваемых при сплайсинге связок с экзонами, поэтому даже если бы ген транскрибировался, интроны бы из РНК не вырезались. Но все перечисленные нарушения перекрывает то, что соответствующие этому псевдогену транскрипты вовсе отсутствуют (вероятно, по причине нарушений в промоторном регионе).

Приведенный список дефектов включает нарушения, потенциально способные помешать каждой из стадий экспрессии гена, и неважно, какое из них изначально инактивировало ген. Если измерить дивергенцию между псевдогеном и функциональным геном, можно оценить период возникновения псевдогена и время начала накопления мутаций.

Если псевдоген потерял активность сразу после дупликации гена β1, следует ожидать равного уровня дивергенции как для молчащих сайтов, так и для сайтов замен (поскольку скорость дивергенции отличается только в случае работы гена за счет создания селективного давления на сайты замен). Для псевдогена Ψβ2 в молчащих сайтах замен оказалось больше. Это означает, что существовал период работы этого псевдогена, в течение которого происходил отбор против изменений в сайтах замен. Основываясь на относительном числе замен в сайтах различных типов, можно определить, что Ψβ2 отделился от β1 около 55 миллионов лет назад, при этом на протяжении 22 миллионов лет он оставался функциональным, а последние 33 миллиона лет существовал в виде псевдогена.

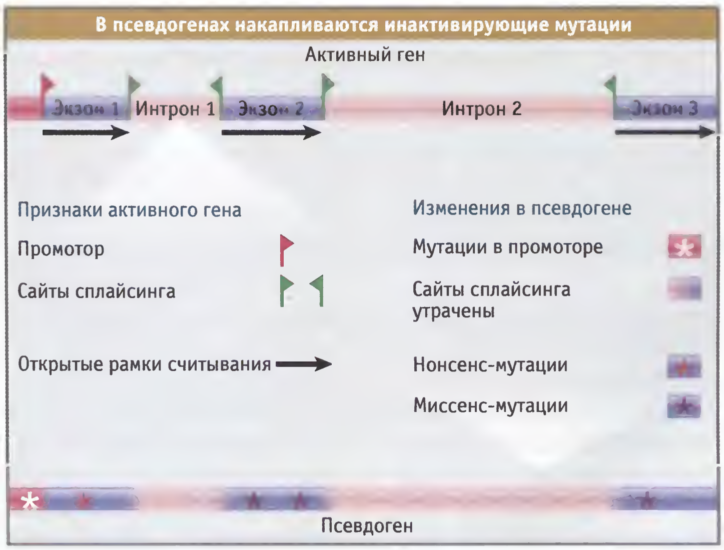


РИС. 6.12 С момента превращения одной из копий β-глобина в псевдоген в ней произошло множество изменений

Похожие расчеты можно провести и для других псевдогенов. Некоторые из них до того как стать псевдогенами, некоторое время были активны, другие же с самого начала были инактивированы. Основной вывод, который можно сделать на основании анализа структуры глобиновых псевдогенов, состоит в том, что на протяжении развития кластеров генов глобинов у разных видов, каждый из них эволюционировал независимо. Это еще раз подтверждает вывод о том, что функциональная дупликация ведет к формированию новых генов, накапливающих отличия и превращающихся либо в новый активный ген, либо в неактивный псевдоген. Она представляет собой длительный процесс, происходящий в пределах кластера генов. Большинство семейств генов содержат псевдогены, однако в целом псевдогены составляют лишь малую часть от общего числа генов в геноме.

Мышиный ген **Ψα**3-глобина обладает интересной особенностью: у него исключительно точно утеряны оба интрона. Последовательность **Ψα**3 можно сопоставить (с учетом накопленных мутаций) с мРНК α-глобина. Его дупликация и инактивация произошли одновременно, из чего можно сделать вывод, что инактивация была связана с потерей интронов.

Неактивные последовательности генома, похожие на РНК-транскрипты, называют **процессированными псевдогенами**. Согласно механизму ретротранспозиции, процессированные псевдогены могут возникать путем вставки в случайное место генома ДНК-копии, полученной с мРНК (см. гл. 22 «Ретровирусы и ретропозоны»). Их характерные особенности приведены на рис. 22.19.

Если псевдогены — это тупики эволюции и очевидно нежелательные последствия изменений функциональных генов, почему они сохраняются в геноме? Выполняют они какую-нибудь функцию или же они бессмысленны и направленный на их сохранение отбор отсутствует?

Необходимо помнить, что мы рассматриваем псевдогены, присутствующие в современных популяциях, однако в ходе эволюции могло быть элиминировано огромное число других псевдогенов. Удаление псевдогена могло происходить путем делеции протяженного участка последовательности или же путем накопления мутаций до уровня, когда псевдоген уже невозможно приписать к какому-либо семейству последовательностей (вероятно, это единственно возможная судьба для псевдогена, который еще не удален).

Удивительно, но даже такие «пережитки эволюции», как псевдогены, могут дуплицироваться. Гены β-глобинов коз представлены тремя зрелыми формами βA, βB и βC (рис. 6.7). Слева от каждой из них есть псевдоген, расположенный в нескольких тпн от активного гена. Псевдогены в большей степени родственны друг другу, нежели зрелым генам β-глобинов (в частности, у них встречается несколько общих инактивирующих мутаций). В свою очередь, последовательности генов β-глобинов больше похожи друг на друга, чем на псевдогены. Это означает, что исходная **Ψβ**-β-структура дуплицировалась и дала начало функциональным β-генам (которые впоследствии приобрели различия) и нефункциональным псевдогенам (превратившимся в современные псевдогены).

Механизмы, ответственные за дупликацию, делецию и перестройку гена, действуют на все последовательности внутри кластера, вне зависимости от того, функциональны они или нет. Окончательная дифференциация продуктов осуществляется путем естественного отбора.

По определению псевдогены не дают белков и обычно не несут никакой функции (за исключением разве что регуляторной). Так, транскрипция псевдогена ингибирует деградацию мРНК, синтезированной с гомологичного ему активного гена. Скорее всего, за деградацию отвечает белок, связывающийся со специфической последовательностью на мРНК. Если эта последовательность есть в РНК, транскрибируемой с псевдогена, эффект, производимый этим белком, может снизиться. Неясно, насколько распространен этот механизм, однако надо помнить, что для транскрибируемых псевдогенов такой тип регуляции возможен.

6.7 Неравный кроссинговер приводит к реорганизации генетических кластеров

Ключевые понятия:

- Если геном содержит кластер генов со схожими последовательностями, спаривание неаллельных генов может приводить к неравному кроссинговеру. Это ведет к формированию делеции в одной из рекомбинантных хромосом и соответствующей дупликации в другой.
- Различные талассемии вызваны делециями, элиминирующими α- или β-глобиновые гены. Тяжесть заболевания у данного индивида зависит от места делеции.

Перестройки внутри кластера похожих или идентичных генов возникают довольно часто. Результат таких перестроек можно увидеть, сравнивая кластеры β-глобинов млекопитающих (рис. 6.7). Несмотря на то что кластеры выполняют схожие функции и имеют общую структуру, они отличаются по длине вследствие разницы в числе и типе β-глобиновых генов, а также числе и строении псевдогенов. Все эти отличия должны были возникнуть с момента появления млекопитающих около 85 млн лет назад (последняя общая для всех млекопитающих точка в эволюции).

Сравнение кластеров показывает, что дупликация, перестройка и вариация генов — такой же важный фактор эволюции, как и медленное накопление точечных мутаций в отдельных генах. Каковы же механизмы, ответственные за реорганизацию гена?

Как было сказано во введении к этой главе, неравный кроссинговер происходит в результате спаривания двух *не*соответствующих сайтов. Обычно

рекомбинация затрагивает одинаковые участки двух гомологичных хромосом. В случае присутствия на каждой из хромосом двух копий гена, между ними может возникать неправильное спаривание (для этого необходимо, чтобы и некоторые из соседних участков оставались неспаренными). Подобные ошибки кроссинговера могут произойти на участке коротких повторов (см. рис. 6.3) или в кластере генов. На **РИС. 6.13** показано, что неравный кроссинговер в кластере генов может приводить к двум следствиям: качественному и количественному:

- Число повторов увеличивается в одной хромосоме и уменьшается в другой (поскольку одна рекомбинантная хромосома приобретает делецию, а другая — инсерцию). Причем результат будет один, вне зависимости от точного расположения перекреста. У одного из изображенных на рисунке рекомбинантов число копий гена увеличивается с двух до трех, а у другого — уменьшается с двух до одного.
- Если рекомбинация происходит в пределах гена (а не между ними), результат будет зависеть от того, идентичны рекомбинирующие гены или только похожи. Если рекомбинирующие копии генов полностью идентичны, то изменений в последовательности каждого из генов не будет. Неравный кроссинговер может происходить и для схожих, но не идентичных генов (хотя в этом случае вероятность возникновения ошибок спаривания, видимо, ниже). В этом случае каждый из рекомбинантных генов будет отличаться от родительских последовательностей.

Получит новая хромосома селективное преимущество или будет испытывать давление отбора — зависит как от последствий изменения последовательности гена, так и от изменения числа копий гена.

Препятствием для неравного кроссинговера может послужить прерывистая структура генов. В случаях, подобных глобинам, соответствующие экзоны прилегающих копий генов, по-видимому, имеют достаточное сходство для того, чтобы спариваться друг с другом, тогда как последовательности интронов могут сильно отличаться. Ограничение спаривания экзонов значительно сокращает длину участка ДНК, вовлекаемого в кроссинговер, а это понижает вероятность неравного кроссинговера. Таким образом, вариабельность интронов повышает стабильность кластеров генов, препятствуя возникновению неравного кроссинговера.

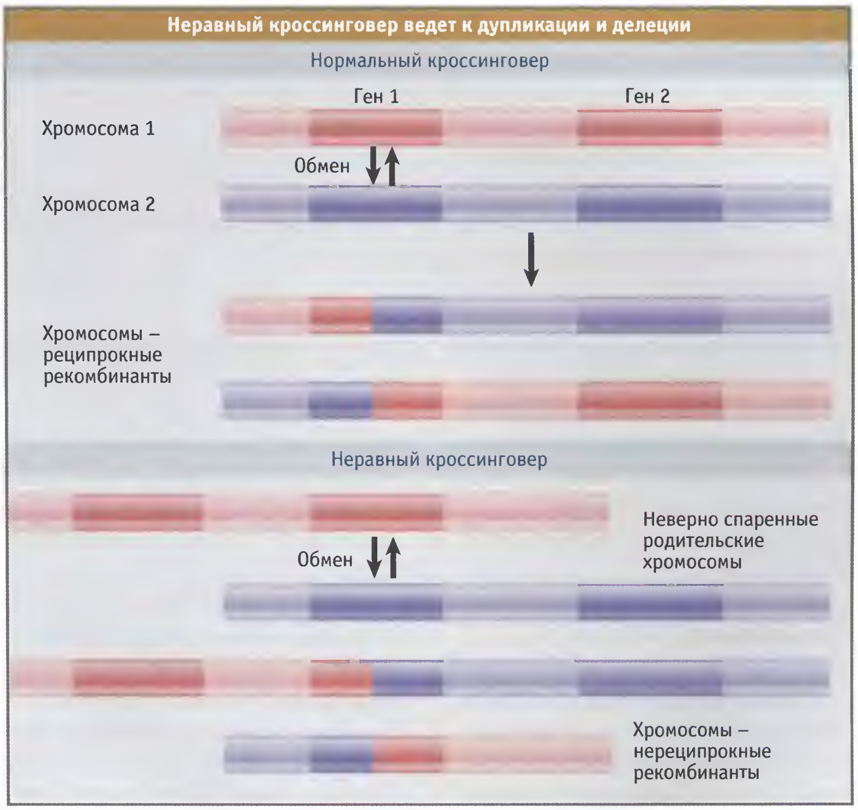


РИС. 6.13 Число генов может измениться в результате неравного кроссинговера. Если ген 1 одной хромосомы образует пару с геном 2 (а не с геном 1) другой хромосомы, вторые копии генов не смогут встать в пару. В результате рекомбинация между неверно спаренными генами порождает хромосому с единственной (рекомбинантной) копией гена и хромосому с тремя копиями гена (по одной от каждого родителя и одну рекомбинантную)

Талассемии возникают в результате мутаций, ослабляющих или прекращающих синтез α- или β-глобина. Природа некоторых талассемий объясняется неравным кроссинговером в кластерах генов глобина человека.

Большинство из наиболее тяжелых талассемий являются результатом делеции части кластера. По крайней мере, в некоторых случаях концы делеций расположены в гомологичных регионах, что очень похоже на ситуацию с неравным кроссинговером.

На РИС. 6.14 приведены делеции, вызывающие α-талассемии. Делеции α-thal-1 затрагивают протяженные участки ДНК и различаются по локализации своих левых концов, тогда как положение правых концов делеций находится за пределами кластера. Эти делеции элиминируют оба гена α. Делеции α-thal-2 сравнительно короткие и элиминируют лишь один из двух генов α. Делеция типа L приводит к удалению участка ДНК в 4,2 тпн, включающего в себя ген α2. Вероятно, эта делеция возникает в результате неравного кроссинговера, поскольку концы делеции расположены в гомологичных регионах хромосомы, находящихся чуть правее гена ψα и чуть правее гена α2. Делеция типа R является результатом удаления фрагмента ДНК размером 3,7 тпн, что в точности соответствует расстоянию между генами α1 и α2. Вероятно, она возникла в результате неравного кроссинговера на участке между этими генами. Описанная ситуация изображена на рис. 6.13.

В зависимости от диплоидной комбинации талассемичных хромосом, пациент может нести любое число α-цепей (от нуля до трех). Между индивидуумами с двумя или тремя α-генами и индивидами дикого типа (четыре α-гена) отличия незначительны. Если же человек несет только один α-ген, то избыток β-цепей приводит к формированию тетрамера β_4, вызывающего заболевание **гемоглобинопатию** Н **(HbH)**. Полное отсутствие α-генов приводит к развитию **водянки плода** и его гибели до или вскоре после рождения.

Неравный кроссинговер (аналогичный тому, который привел к формированию талассемичных хромосом) может привести к формированию хромосомы с тремя α-генами. Люди, несущие такие хромосомы, были обнаружены в нескольких популяциях. В некоторых популяциях частота встречаемости тройного α-локуса примерно равна частоте встречаемости одиночного α-локуса; в других популяциях тройные α-гены встречаются значительно *реже*, чем одиночные. Это означает, что на различные популяции действовали (неизвестные) факторы отбора, поддерживающие определенный уровень генов глобинов.

Вариации числа α-генов встречаются довольно часто, означая, что неравный кроссинговер в кластере весьма распространен. Создается впечатление, что в α-кластере неравный кроссинговер происходит чаще, чем в β-кластере. Одна из возможных причин этого состоит в том, что интроны α-генов короче и поэтому в меньшей степени препятствуют неправильному спариванию между негомологичными генами.

Делеции, вызывающие β-талассемии, приведены на РИС. 6.15. В некоторых (довольно редких) случаях они затрагивают только ген β. В этих случаях обнаруживается делеция в 600 пн, затрагивающая участок от второго интрона вплоть до 3'-фланкирующего региона. В других случаях делеции затрагивают сразу несколько генов. Могие делеции довольно протяженны и простираются на >50 тпн от 5'-конца.

Гемоглобин типа **Hb Lepore** является классическим примером возникновения делеции в результате неравного кроссинговера между сцепленными генами. Последовательности β- и δ-генов отличаются всего на ~7%. Неравная рекомбинация приводит к удалению участка ДНК, расположенного между этими генами, соединяя их (рис. 6.13). Объединенный ген продуцирует одну β-подобную цепь, состоящую из N-концевой последовательности δ-цепи и C-концевой последовательности β-цепи.

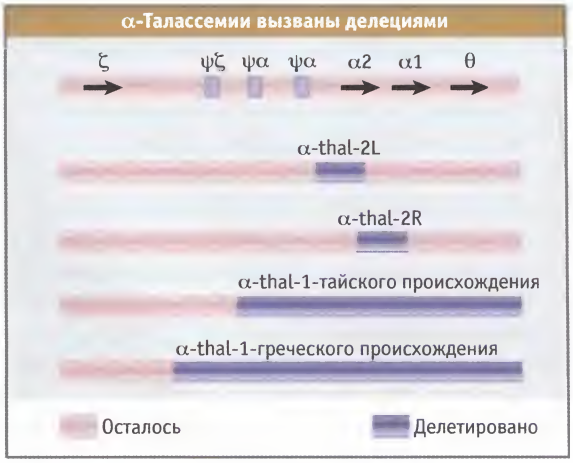


РИС. 6.14 Различные формы α-талассемии развиваются из-за различных делеций в α-глобиновом кластере

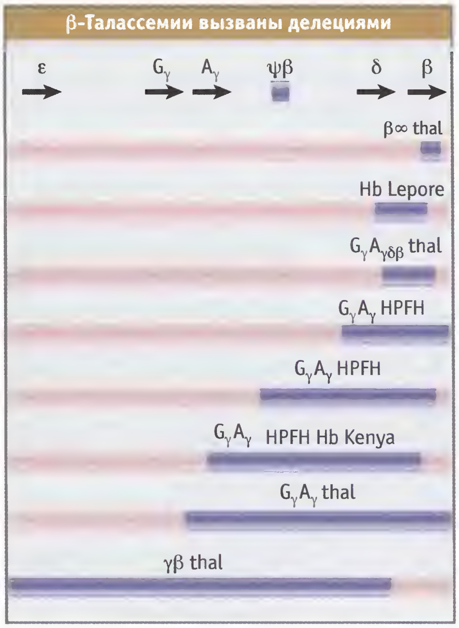


РИС. 6.15 Делеции в β-глобиновом кластере являются причиной нескольких форм талассемии

На сегодняшний день известно несколько типов гемоглобинов Lepore, отличающихся местом перехода от δ- к β-последовательности. Поэтому при возникновении неравного кроссинговера между спаривающимися δ- и β-генами точная точка рекомбинации определяет место в аминокислотной цепи, где произойдет переключение с δ- на β-последовательность.

Явление, обратное описанному, приводит к возникновению гемоглобина типа **анти-Lepore**, продуцируемого геном, кодирующим N-концевую часть β- и С-концевую часть δ-последовательности. Такой «химерный» ген располагается между нормальными δ- и β-генами.

В пользу того, что неравный кроссинговер может происходить и между «дальними родственниками», свидетельствует обнаружение еще одного «химерного» гемоглобина — **Kenya**. Он содержит N-концевую последовательность гена $^{A}\gamma$ и С-концевую последовательность гена β. Эта «химера», скорее всего, также является результатом неравного кроссинговера между $^{A}\gamma$ и β, последовательности которых отличаются примерно на 20%.

На основе различий между кластерами глобиновых генов у различных млекопитающих можно сделать вывод о том, что дупликация и (иногда) следующее за ней накопление отличий сыграли важную роль в эволюции каждого кластера. Талассемические делеции человека демонстрируют, что неравный кроссинговер по-прежнему возникает в обоих кластерах генов глобина. Каждое такое событие приводит к возникновению как дупликации, так и делеции, и мы можем проследить судьбу каждого из двух рекомбинантных локусов в популяции. Делеции (в принципе) могут возникать и в результате рекомбинации между гомологичными последовательностями, расположенными *на одной и той же* хромосоме, причем такой вариант не влечет за собой появление соответствующей дупликации.

Довольно сложно оценить естественную частоту этих процессов, поскольку давление отбора быстро выравнивает уровни вариабельных кластеров в популяции. Скорее всего, отбор в основном работает против уменьшения числа генов (в результате делеций). В некоторых популяциях может поддерживаться баланс, допускающий существование незначительной доли хромосом, несущих делецию и обладающих в связи с этим неким преимуществом.

Структуры современных кластеров человека обнаруживают несколько дупликаций, подтверждающих значимость описанных механизмов. *Функциональные* последовательности включают два гена α, кодирующих один и тот же белок, довольно близкородственные гены β и δ, а также два почти идентичных гена γ. Эти сравнительно недавние независимые дупликации сохранились в популяции безотносительно к более древним дупликациям, давшим начало различным типам генов глобинов. Другие дупликации могли способствовать возникновению псевдогенов или могли быть утеряны. Очевидно, что дупликации и делеции играют важную роль в формировании всех генных кластеров.

6.8 Гены рРНК формируют тандемные повторы

Ключевые понятия:

- Рибосомальная РНК кодируется большим числом идентичных тандемно повторяющихся генов, формирующих один или несколько кластеров.
- Каждый кластер рДНК организован таким образом, что кодирующие предшественник рРНК транскрипционные единицы перемежаются нетранскрибируемыми спейсерами.

В примерах, которые мы обсуждали до сих пор, между отдельными членами кластера генов существовали различия, позволявшие отбору оказывать независимое влияние на каждый ген. Обратная ситуация наблюдается для двух крупных кластеров генов, содержащих множество идентичных копий одного и того же гена (или генов). Большинство организмов содержат множественные копии генов, кодирующих гистоны, — белки, являющиеся важным компонентом хромосом. Кроме того, почти все организмы несут множественные копии генов, кодирующих рибосомальную РНК. Эти случаи ставят перед нами интересные эволюционные вопросы.

Рибосомальная РНК — это основной продукт транскрипции, составляющий как у прокариот, так и у эукариот около 80–90% от общей массы клеточной РНК. Число генов рРНК варьирует от семи у *E. coli* до 100–200 у низших эукариот и нескольких сотен у высших эукариот. Гены, кодирующие большую и малую рРНК (входящие в состав, соответственно, большой и малой субъединиц рибосомы), обычно формируют тандемную пару (единственным исключением являются гены рРНК митохондриального генома дрожжей).

Отсутствие выраженных различий в последовательностях молекул рРНК свидетельствует о том, что все копии каждого из генов должны быть идентичны (или, по крайней мере, их отличия должны быть на уровне ниже детектируемого (~1%)). Следовательно, должен существовать механизм (или механизмы), препятствующий накоплению вариаций в отдельных последовательностях.

У бактерий множественные пары генов рРНК рассредоточены по геному. В большинстве же геномов эукариот гены рРНК собраны в тандемный кластер (или кластеры). Иногда эти участки называют **рДНК**. (В некоторых случаях большое количество рДНК и ее необычный нуклеотидный состав позволяют выделить эту ДНК в виде отдельной фракции из фрагментированной геномной ДНК.) Важным признаком тандемных кластеров является кольцевая карта их рестрикции (как показано на РИС. 6.16).

Предположим, что каждый отдельно взятый повтор содержит три сайта рестрикции. Если нанести эти фрагменты на обычную рестрикционную карту, мы обнаружим, что за фрагментом A следует фрагмент B, за которым следует C, за которым расположен A, и т. д. Таким образом, фрагменты формируют как бы

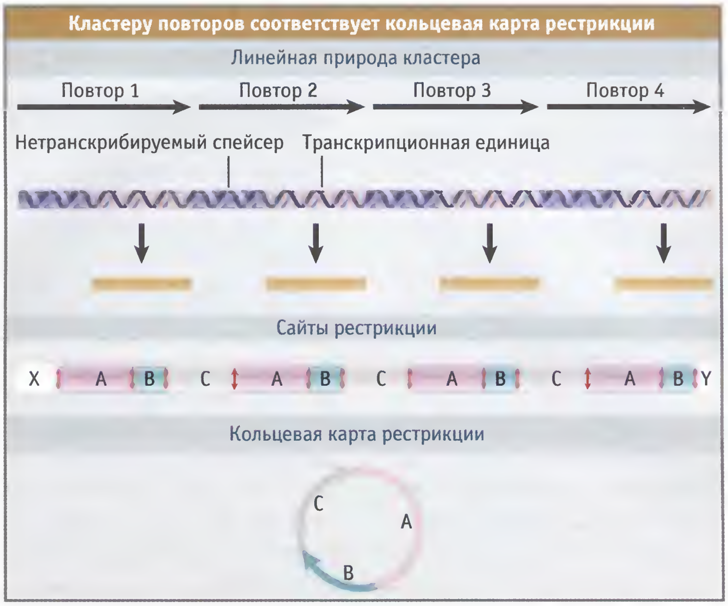


РИС. 6.16 В кластере тандемных копий гена транскрипционные единицы чередуются с нетранскрибируемыми спейсерами; карту рестрикции такого кластера можно представить в виде кольца

кольцевую карту (как если бы ДНК действительно была кольцевой). Если кластер большой, то внутренние фрагменты (A, B и C) будут представлены в гораздо большем количестве, чем концевые фрагменты (X и Y), соединяющие кластер с соседними участками ДНК. В кластере, состоящем из 100 повторов, на долю X и Y будет приходиться 1% от уровня A, B, и C. Это может усложнить получение концов кластера генов и, как следствие, помешать при создании карты.

В ядре имеется особый регион, в котором происходит синтез рРНК — **ядрышко**. Фибриллярная центральная часть ядрышка окружена гранулярным кортексом. В фибриллярном центре осуществляется транскрипция рРНК на матрице ДНК, а гранулярный кортекс сформирован частицами рибонуклеопротеина, в которые встраивается рРНК. Особенности морфологии ядрышка продемонстрированы на РИС. 6.17.

Отдельные участки хромосом, пространственно ассоциированные с ядрышком, называют **ядрышковыми организаторами**. Каждый ядрышковый организатор представляет собой кластер тандемно повторяющихся генов рРНК, расположенных на одной хромосоме. Сосредоточенность тандемно повторяющихся генов рРНК в определенном месте ядра, а также интенсивность их транскрипции, обусловливает характерную морфологию ядрышка.

Пара основных рРНК транскрибируется в бактериальных клетках и эукариотических ядрах в виде единого предшественника. После транскрипции предшественник разрезается на две отдельные молекулы рРНК. Самый короткий рРНК-предшественник — у бактерий, а самый длинный — у млекопитающих (у которых его, в соответствии со скоростью

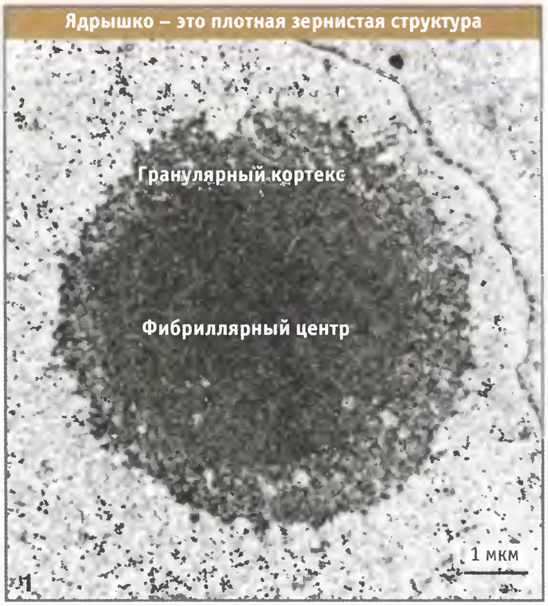


РИС. 6.17 На тонком срезе клетки тритона *Notophtalmus viridescens* хорошо различимо ядрышко. Волокнистая (фибриллярная) сердцевина ядрышка состоит из транскрибируемой рДНК; сердцевину ядрышка окружает зернистый (гранулярный) кортекс, состоящий из рибосомальных субъединиц в состоянии сборки. Фотография любезно предоставлена Oscar Miller

осаждения при центрифугировании (седиментации), называют 45S РНК). Кластер рДНК содержит множество единиц транскрипции, каждая из которых отделена от предыдущей **нетранскрибируемым спейсером**. Различия между транскрипционной единицей



РИС. 6.18 Транскрипция кластеров рДНК выглядит как серия транскрипционных единиц. Каждая такая единица («елочка») отделена от предыдущей и последующей нетранскрибируемыми спейсерами. Фотография любезно предоставлена Oscar Miller

и нетранскрибируемым спейсером можно увидеть на фотографиях, полученных с помощью электронного микроскопа. Препарат, показанный на РИС. 6.18, получен из клеток тритона *Notopthalmus viridescens*. На фотографии видно, что синтез РНК на каждой транскрипционной единице происходит настолько интенсивно, что в транскрипцию одной повторяющейся единицы одновременно осуществляет множество РНК-полимераз. Полимеразы расположены столь плотно, что РНК-транскрипты образуют характерную форму, отражающую увеличение длины продукта при продвижении фермента по транскрипционной единице (от промотора к терминатору).

6.9 Множественные гены рРНК сохраняют постоянство своей последовательности

Ключевые понятия:

- Все гены, расположенные в кластере рДНК, идентичны по своей последовательности.
- Каждый нетранскрибируемый спейсер состоит из более коротких повторяющихся единиц, число которых варьирует таким образом, что длины отдельных спейсеров различаются.

Нетранскрибируемые спейсеры рДНК сильно варьируют по длине как между предствителями различных видов, так и (иногда) в пределах вида. Короткий нетранскрибируемый спейсер дрожжей имеет относительно постоянную длину. У *D.melanogaster* между различными копиями транскрипционных единиц наблюдается почти двукратная вариация длины нетранскрибируемого спейсера. Похожая ситуация и у *X. laevis*. В каждом из этих случаев все транскрипционные единицы представлены одним тандемным кластером, расположенным на определенной хромосоме. (В случае с *D.melanogaster* — это половая хромосома. Кластер на X хромосоме больше, чем на Y, за счет чего у самок *D.melanogaster* больше копий генов рРНК, чем у самцов.)

У млекопитающих повторяющаяся единица гораздо длиннее, чем у насекомых и земноводных. Она включает транскрипционную единицу размером ~13 тпн и нетранскрибируемый спейсер в ~30 тпн. Обычно гены сгруппированы в несколько рассредоточенных по разным хромосомам кластеров — у человека и мыши кластеры находятся на пятой и шестой хромосомах соответственно.

Как же корректирующие механизмы, действующие, по всей видимости, в пределах одного кластера и поддерживающие постоянство последовательности рРНК, работают при наличии нескольких кластеров? Это интересный вопрос, но ответа на него пока нет.

Изменчивость длины нетранскрибируемых спейсеров, принадлежащих одному кластеру генов, контрастирует с консервативностью последовательности транскрипционной единицы. Но даже несмотря на эту изменчивость, последовательности более длинных нетранскрибируемых спейсеров гомологичны коротким спейсерам. Дело в том, что каждый нетранскрибируемый спейсер содержит ***внутренние повторы***, и изменчивость его длины, таким образом, является результатом изменения числа повторов некоторой «элементарной» субъединицы.

На РИС. 6.19 изображена общая структура нетранскрибируемого спейсера *X. laevis*. Участки фиксированной длины чередуются с варьирующими по длине регионами. Каждый из трех повторяющихся участков состоит из различного числа более коротких повторов. Один из участков содержит повторы длиной 97 пн; два других содержат повторяющиеся единицы длиной 60 пн и 80 пн. Вариация числа таких единиц в регионах, содержащих повторы, вносит свой вклад в изменчивость всего спейсера. Регионы, содержащие повторы, разделены так называемыми **Bam-островками** — высококонсервативными короткими участками (Bam-островки получили свое название благодаря их обнаружению при помощи рестриктазы BamHI). Описанный тип организации свидетельствует о том, что кластер эволюционировал путем дупликаций, включая и промоторный регион.

Попробуем разобраться, почему же последовательности экспрессируемых копий повторяющихся генов так незначительно отличаются друг от друга. В соответствии с одной из моделей у клетки существует потребность в определенном числе «хороших» последовательностей. Однако эта модель не запрещает мутировавшим последовательностям накапливаться вплоть до того момента, когда их доля в кластере станет достаточной для запуска процессов селективного отбора. Поскольку вариаций такого типа в кластерах явно недостаточно, подобные модели можно исключить из рассмотрения.

Небольшое число вариаций последовательности генов указывает на наличие отбора, чувствительно-

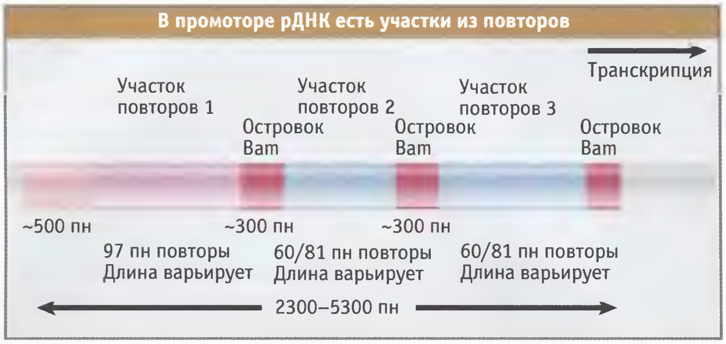


РИС. 6.19 Нетранскрибируемый спейсер рДНК *X. laevis* имеет внутреннюю ритмическую структуру из повторов. Число этих повторов (и, как следствие, длина участков, состоящих из этих повторов) варьирует. Участки повторов разделены островками Bam — небольшими сегментами ДНК с одной и той же последовательностью

го к индивидуальным изменениям в гене. Еще одна модель основана на предположении о периодическом обновлении всего кластера за счет одного или нескольких его представителей. В принципе, любой подобный механизм должен включать обновление в каждом поколении. Однако эту модель тоже можно исключить, поскольку она предполагает отсутствие различий и в нетранскрибируемых участках отдельных повторов.

Дилемма до сих пор остается неразрешенной. Изменчивость нетранскрибируемых регионов указывает на то, что по ним довольно часто происходит неравный кроссинговер. Это ведет к изменению размера кластера, но последовательности отдельных повторов остаются неизменны. Так что же создает помехи для накопления мутаций? В следующем разделе мы увидим, что приведение к одному виду (гомогенизация) копий кластера является результатом его постоянного расширения и сужения.

6.10 Рекомбинационная фиксация может поддерживать идентичность повторов

Ключевые понятия:

- Неравный кроссинговер изменяет размер кластеров тандемных повторов.
- Отдельные повторяющиеся единицы могут быть элиминированы из кластера или вынесены за его пределы.

При рассмотрении активных дуплицированных генов возникает следующий вопрос: как отбор предотвращает накопление мутаций в «лишних» копиях гена?

Теоретически дупликация гена приводит к немедленному ослаблению эволюционного давления на его последовательность. Теперь, когда в геноме есть две идентичные копии, изменение последовательности любой из них не лишит организм функционального белка, поскольку исходную аминокислотную последовательность продолжает кодировать другая копия. Давление отбора распределится на обе копии до тех пор, пока исходная функция одной из них не изменится настолько, что отбор начнет действовать на них независимо.

Сразу после дупликации гена изменения в одной из копий должны накапливаться быстрее, что, в конце концов, приведет к возникновению новой функции (или к потере функции и формированию псевдогена). Если у гена появляется новая функция, то он продолжает эволюционировать, но уже медленнее, приблизительно со скоростью, характерной для эволюции исходного гена. Вероятно, подобные процессы ответственны за разделение функций между генами эмбрионального и зрелого глобина.

Однако существуют примеры, когда дуплицированные гены сохраняют свою исходную функцию и продолжают кодировать идентичные (или почти идентичные) белки. Так, у человека идентичные белки кодируют два гена α-глобина, а два белка γ-глобина отличаются лишь одной аминокислотой. Каким образом селективное давление помогло сохранить идентичность этих последовательностей?

Вероятнее всего, функции двух генов (в каждом из приведенных случаев) отличаются неким (неизвестным) свойством, например местом или временем экспрессии. Другая гипотеза состоит в том, что две копии нужны для обеспечения продукции определенного количества белка, и ни одна из них сама по себе не справляется с производством его необходимого объема.

Однако в ряде случаев, когда повторяемость гена особенно велика, нельзя исключить, что хотя бы одна из множества копий не является жизненно важной. Если у гена много активных копий, то эффект от возникновения мутации в любой из них

должен быть очень незначительным. Последствия от возникновения отдельных мутаций нивелируются большим числом активных копий, сохраняющих последовательность дикого типа. Прежде чем на организм будет оказан летальный эффект очередной мутации, в геноме должно накопиться много мутантных копий.

Для множественных генов летальность становится «количественной», что можно продемонстрировать на примере кластеров рДНК *X. laevis* или *D. melanogaster*. Из этих кластеров можно удалить половину повторяющихся единиц без заметного влияния на фенотип. Так как же эти единицы избежали постепенного накопления вредных мутаций? Какова вероятность того, что возникшая полезная мутация сможет продемонстрировать свои преимущества в кластере?

Основной принцип моделей, объясняющих поддержание идентичности повторяющихся копий, состоит в предположении, что повторяющиеся гены наследуются не независимо, а могут вновь и вновь воспроизводиться с *одной* из копий. Простейшим примером могут служить два идентичных гена. При возникновении мутации в одной из копий она либо элиминируется (за счет переноса последовательности второй копии), либо распространяется на обе копии (если новая версия оказывается более выгодной). В обоих случаях мутация служит материалом для естественного отбора. В результате два гена эволюционируют вместе, как будто они представляют собой единый локус. Такое явление называют **согласованной**, или **сопряженной эволюцией** (иногда **коэволюцией**).

Один из механизмов предполагает, что последовательности повторяющихся генов непосредственно сравниваются друг с другом и гомогенизируются под действием особых ферментов, распознающих любые различия в последовательности ДНК. Например, гомогенизация может происходить путем обмена одиночными цепями ДНК, в результате чего формируются гены, у которых первая цепь происходит от одной копии, а вторая — от другой. В этом случае любые различия между ними проявляются в виде неспаренных оснований, что привлекает внимание ферментов репарации, удаляющих несоответствия так, чтобы оставались только пары A-T и G-C. Этот процесс связан с рекомбинацией и называется **генной конверсией**. Определить, имеет ли этот процесс место в конкретном случае, можно путем сравнения последовательностей дуплицированных генов. Если они эволюционируют согласованно, мы не должны увидеть накопление замен в молчащих сайтах (поскольку гомогенизация затрагивает их в той же степени, что и сайты замещения). Известно, что механизм восстановления необязательно распространяется за пределы гена, поскольку в некоторых случаях дуплицированные гены имеют совершенно разные фланкирующие последовательности. Действительно, можно заметить, что границы гомогенизированных последовательностей резко выделяются.

Необходимо помнить, что наличие описанного механизма может свести на нет точность оценки периода, прошедшего с момента дупликации таких генов, по их дивергенции, *поскольку дивергенция в данном случае отражает время, прошедшее не с момента исходной дупликации, а с момента последнего «выравнивания»*.

Другая модель «выравнивания» получила название **кроссоверной фиксации**. Согласно этой модели, весь кластер служит предметом непрерывной реорганизации по механизму неравного кроссинговера. Если неравный кроссинговер приводит к физическому воспроизведению всех копий из одной, то процессы, описываемые моделью перекрестного закрепления, могут объяснить сопряженную эволюцию генов.

Продолжая пример, изображенный на рис. 6.13, можно представить, что хромосома, несущая три копии гена, может потерять один ген в результате делеции. Тогда последовательности двух оставшихся генов будут на $1^1/_2$ состоять из последовательности одной из исходных копий и на $^1/_2$ из последовательности второй копии гена. Мутация, расположенная в первом (более длинном) регионе, теперь может затрагивать оба гена и начнет испытывать на себе давление естественного отбора.

Тандемное расположение генов в кластерах открывает широкие возможности для неправильного спаривания генов, последовательности которых одинаковы, но расположены в различных местах кластера. За счет неравного кроссинговера, приводящего к непрерывному увеличению и уменьшению числа единиц, составляющих кластеры, может возникать ситуация, при которой все единицы протяженного кластера берут начало всего от нескольких единиц предкового кластера. Различная длина спейсеров согласуется с идеей о том, что неравный кроссинговер происходит по спейсерам и может объяснить гомогенность генов в сравнении с вариабельностью спейсеров. Благодаря механизму «выравнивания», составляющие кластер множественные гены подвержены давлению отбора; однако на спейсеры отбор не действует, и поэтому они сохраняют способность накапливать отличия.

На участках ДНК с уникальными последовательностями рекомбинация возникает между соответствующими точками на гомологичных хромосомах, давая, таким образом, реципрокных рекомбинантов. В основе такой точности лежит способность двух дуплексов ДНК к специфичному спариванию. Известно, что неравная рекомбинация может происходить даже в том случае, если гены гомологичны лишь своими экзонами (несмотря на то что фланкирующие и промежуточные участки могут различаться). Это происходит из-за неправильного спаривания между соответствующими экзонами *неаллельных* генов.

Представьте, насколько чаще должно возникать неправильное спаривание в тандемном кластере идентичных или почти идентичных повторов. За исключением концевых участков кластера, высокое сходство повторов делает практически невозможной идентификацию точно соответствующего повтора! Это имеет два последствия: непрерывную регулировку размера кластера и гомогенизацию повторяющихся единиц.

Рассмотрим последовательность, состоящую из повторяющейся единицы «ab» с концевыми участками «x» и «y». Если последовательность одной хромосомы покрасить черным, а последовательность другой красным, то точное выравнивание между «аллельными» последовательностями будет выглядеть следующим образом:

xababababababababababababababababy
xababababababababababababababababy

Однако для *любой* единицы *ab*, расположенной на одной хромосоме, существует вероятность спаривания с *любой* другой единицей *ab*, расположенной на другой хромосоме. Неправильное выравнивание может, например, выглядеть так:

xababababababababababababababababy
    xababababababababababababababababy

Участок спаривания в этом случае ничуть не менее стабилен, чем идеально выровненная пара, просто он немного короче. Процесс спаривания хромосом перед рекомбинацией изучен недостаточно, однако, вероятнее всего, он начинается с взаимодействия коротких участков, распространяясь затем на всю длину хромосомы. Если процесс спаривания начнется в пределах сателлитной ДНК, то он скорее произойдет со сдвигом, чем в точном соответствии с местоположением повторяющегося элемента в кластере.

Теперь предположим, что рекомбинация продолжается в пределах неправильно спаренного участка. Рекомбинанты в таком случае будут содержать различное число повторяющихся единиц. Один из кластеров станет длиннее, а другой короче,

xabababababababababababababababababy
xababababababababax
  xabababababababababababababababababababy
↓
xabababababababababbabababababababababababy
+
xabababababababaabababababababababy

где «x» обозначает место перекреста.

Обычно процессы подобного рода приводят к постоянному увеличению и уменьшению длины кластеров тандемных повторов. Процесс может привести к тому, что определенная повторяющаяся единица займет весь кластер (РИС. 6.20). Предположим, что изначально кластер состоял из последовательности *abcde*, где каждая буква обозначает определенную повторяющуюся единицу. Различные повторяющиеся единицы связаны друг с другом родством, достаточным для того, чтобы допустить неправильное спаривание при рекомбинации. Затем, в результате серии неравных рекомбинаций, размер повторяющегося участка увеличивался и уменьшался, в результате чего одна из единиц получила преимущественное распространение, заменив остальные.

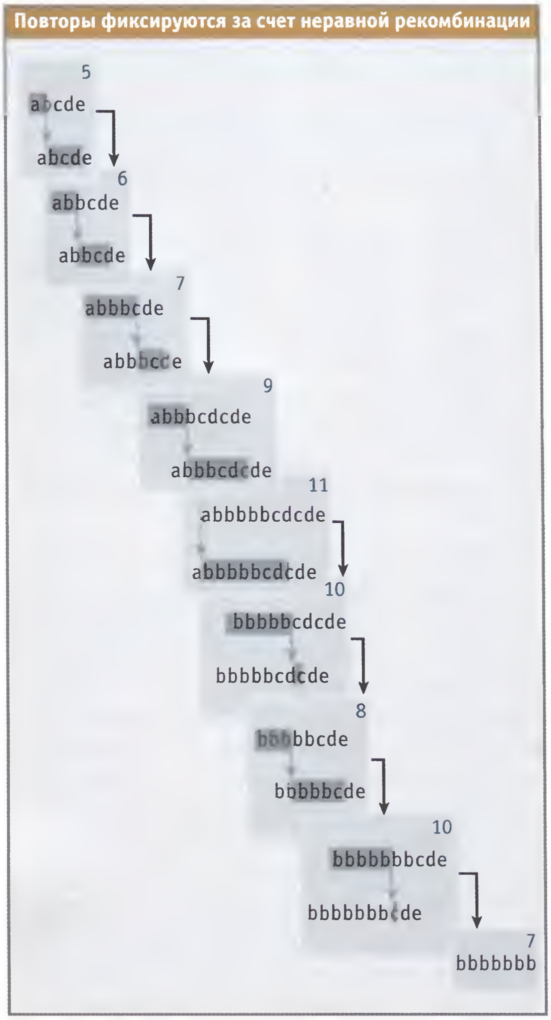


РИС. 6.20 Неравная рекомбинация позволяет одной конкретной повторяющейся единице геномной последовательности занять целый кластер. Числа обозначают длину кластера на разных этапах экспансии последовательности b

Модель кроссоверной фиксации гласит, что *любая последовательность ДНК, не испытывающая на себе давления отбора, постепенно превратится в серию идентичных тандемных повторов за счет процессов, описываемых этой моделью.* Важным допущением при этом является то, что процесс фиксации

происходит быстрее процесса накопления мутаций, и, таким образом, новые мутации либо элиминируются (в результате удаления несущих их повторов), либо распространяются на весь кластер. В случае с кластером рДНК его дальнейшее развитие обусловлено отбором эффективно транскрибируемых последовательностей.

6.11 Сателлитная ДНК часто расположена в гетерохроматине

Ключевые понятия:

- Высокоповторяющаяся ДНК состоит из очень коротких повторяющихся последовательностей и не несет кодирующей функции.
- Она встречается в виде крупных блоков, которые могут иметь различные физические свойства.
- Она часто является основной составной частью центромерного гетерохроматина.

Характерной особенностью участков повторяющейся ДНК является их сравнительно высокая скорость ренатурации. Быстрее всех ренатурирует компонент эукариотических геномов, называемый высокоповторяющейся ДНК, который состоит из коротких, многократно повторенных последовательностей, собранных в крупные тандемные кластеры. Поскольку его повторяющиеся единицы довольно коротки, такие участки иногда называют **простой последовательностью ДНК**. Элементы такого типа есть в геномах всех высших эукариот, однако приходящаяся на их долю часть генома сильно варьирует. В геномах млекопитающих они обычно составляют менее 10% последовательности, но, например у *Drosophila virilis* высокоповторяющаяся ДНК занимает около 50% генома. Помимо крупных кластеров, в которых впервые были обнаружены такие последовательности, существуют и меньшие по размеру кластеры, включенные в состав уникальной ДНК. Обычно они состоят из коротких последовательностей, повторяющихся в виде идентичных или близкородственных копий.

Короткие тандемные повторы обычно формируют фракцию с характерными физическими свойствами, позволяющими отделить ее от остальной ДНК. В ряде случаев нуклеотидный состав высокоповторяющихся последовательностей отличается от среднего нуклеотидного состава генома, что позволяет выделить их в отдельную фракцию на основании их характерной плотности. Фракции подобного рода получили название сателлитной ДНК. Термин «сателлитная ДНК» по сути, является синонимом термина «простая последовательность ДНК». Будучи составлена из коротких тандемных повторов, такая ДНК не транскрибируется и не транслируется.

Тандемно повторяющиеся последовательности более других подвержены несовпадениям при спаривании гомологичных хромосом. Поэтому размеры тандемных кластеров очень разнообразны в популяции и, как правило, отличаются на индивидуальном уровне. Небольшие кластеры таких последовательностей можно использовать для характеристики индивидуальных геномов. Метод, использующий этот принцип, называют методом «генетических отпечатков пальцев» (см. разд. 6.14 «Мини-сателлиты можно использовать при генетическом картировании»).

Плотность ДНК-дуплекса в растворе зависит от его G-C состава и выражается эмпирической формулой:

$$\rho = 1{,}660 + 0{,}00098\ (\%\text{G-C})\ \text{г/см}^{-3}$$

Разделение ДНК по плотности удобно проводить методом центрифугирования в **градиенте плотности** CsCl, в котором ДНК формирует полосу в месте, соответствующем ее собственной плотности. В градиенте плотности обычно различимы фракции ДНК, отличающиеся по G-C составу более чем на 5%.

Центрифугируя ДНК в градиенте плотности, можно выявить два компонента генома:

- Большая часть генома состоит из фрагментов, выявляющихся на графике в виде довольно широкого пика, центральная часть которого отражает значение плотности соответствующее среднему G-C составу генома. Это так называемая основная полоса.
- Иногда выявляется меньший по размеру пик (или пики), имеющий другое значение плотности. Это и есть сателлитная ДНК.

Сателлитная ДНК присутствует в большинстве эукариотических геномов. Она может быть как легче, так и тяжелее основной полосы, однако редко составляет более 5% от общей массы ДНК. Хорошим примером является ДНК мышей, результат разделения которой в градиенте плотности изображен на РИС. 6.21. Этот график представляет собой количественное отражение областей, сформированных фрагментированной мышиной ДНК в результате центрифугирования в градиенте плотности CsCl. Основная полоса включает 92% материала генома, и центральная часть ее пика соответствует плотности 1,701 г/см$^{-3}$ (что соответствует среднему G-C составу ~42%, типичному для млекопитающих). Малый пик отражает оставшиеся 8% генома и имеет плотность 1,690 г/см$^{-3}$. Он содержит сателлитную ДНК мышей, G-C состав которой (30%) значительно ниже, чем в среднем по геному.

Расположение сателлитной ДНК при ценрифугировании в градиенте плотности часто аномально. Поэтому при определении нуклеотидного состава сателлита другими методами, нередко оказывается, что он отличается от состава, предсказанного на основании его плавучей плотности. Причина в том, что величина ρ зависит не только от нуклеотидного состава, но и от того, какие пары являются ближайшими соседями.

В большей степени влияние этого фактора будет касаться регулярных последовательностей. Для них отклонение плотностной формулы (не учитывающей этот параметр) будет больше. Кроме того, сателлитная ДНК может быть метилирована, что также оказывает влияние на ее плавучую плотность.

Основная часть высокоповторяющейся ДНК генома может быть выделена в виде сателлитов. Если компонент, состоящий из высокоповторяющейся ДНК, не отделяется при центрифугировании в виде сателлита, то после его выделения другими способами он в конечном итоге оказывается сателлитной ДНК. В целом можно сказать, что высокоповторяющаяся ДНК состоит из множества тандемных повторов, аномально ведущих себя при центрифугировании. Не выявляющиеся при центрифугировании сателлиты называют **скрытыми (криптическими)**. Скрытая и явная сателлитные ДНК вместе обычно содержат все крупные тандемно повторяющиеся блоки высокоповторяющейся ДНК. Если в геноме имеется более одного типа высокоповторяющейся ДНК, каждый из них представлен отдельным сателлитным пиком (хотя иногда разные сателлиты представлены одним пиком).

Где же в геноме расположены блоки высокоповторяющейся ДНК? Развитие техники гибридизации нуклеиновых кислот позволило определить расположение регионов сателлитной ДНК прямо на хромосоме. При **гибридизации *in situ*** хромосомную ДНК денатурируют, обрабатывая распластанные на покровном стекле клетки. Затем к препарату добавляют радиоактивно меченную пробу ДНК или РНК. Проба гибридизуется с комплементарными ей участками денатурированной хромосомы. Расположение участков гибридизации можно определить с помощью авторадиографии (см. рис. 28.19).

Сателлитная ДНК часто обнаруживается в области **гетерохроматина**. Гетерохроматин — термин, используемый для обозначения неактивных участков хромосом, постоянно находящихся в компактной спирализованной форме, в отличие от **эухроматина**, составляющего основную часть генома (см. разд. 28.7 «Хроматин делится на эухроматин и гетерохроматин»). Гетерохроматин обычно расположен в регионе центромер (участков, где при митозе и мейозе формируются кинетохоры, контролирующие продвижение хромосом к полюсам веретена деления). Расположение сателлитной ДНК в прицентромерном регионе может означать, что в хромосоме она выполняет структурную функцию. Эта функция может быть связана с процессом расхождения хромосом при делении клетки.

На РИС. 6.22 приведен пример локализации сателлитной ДНК в хромосомах клеток мыши. В этом случае мечеными оказываются концы хромосом, поскольку именно там расположены центромеры хромосом *Mus musculus*.



РИС. 6.21 При центрифугировании в градиенте плотности хлорида цезия (CsCl) ДНК мыши разделяется на две фракции: основную и сателлитную



РИС. 6.22 Гибридизация радиоактивной пробы с цитологическими препаратами хромосом показала, что сателлитная ДНК мыши находится в центромерах. Фотография любезно предоставлена Mary Lou Pardue and Joseph G. Gall, Carnegie Institution

6.12 Сателлитная ДНК членистоногих состоит из очень коротких идентичных повторов

Ключевые понятия:

- Повторяющиеся единицы сателлитной ДНК членистоногих составляют всего несколько нуклеотидов в длину. Большинство копий последовательности идентичны.

Сателлитная ДНК членистоногих, типичными представителями которых являются насекомые и рако-

образные, по-видимому, довольно гомогенна по составу. Обычно более 90% сателлита составляет одна очень короткая повторяющаяся единица, что существенно упрощает определение его последовательности.

У *Drosophila virilis* есть три основных типа сателлитов и один скрытый сателлит, которые вместе составляют более 40% генома. Последовательности сателлитов приведены на РИС. 6.23. Три основных сателлита состоят из схожих последовательностей. Достаточно заменить один нуклеотид, чтобы сателлитная ДНК I типа превратилась в сателлитную ДНК II или III типа.

Последовательность сателлитной ДНК I типа можно встретить и у других видов *Drosophila*, родственных *D. virilis*, поэтому она, вероятно, появилась у них еще до того, как они разделились на отдельные виды. Последовательности типов II и III, по всей видимости, специфичны для *D. virilis* (то есть они могли образоваться из сателлита I типа уже после разделения).

Главной отличительной чертой этих сателлитов является малая длина повторяющейся единицы: всего 7 пн. Похожие сателлиты встречаются у других видов. Например, в геноме *D. melanogaster* имеется множество сателлитных ДНК, причем некоторые из них имеют очень короткие повторяющиеся единицы (5, 7, 10 или 12 пн). Похожие сателлиты обнаружены у ракообразных.

Сходство последовательностей разных сателлитных ДНК внутри вида (как у *D. virilis*) не является правилом. У других организмов сателлитные ДНК могут существенно различаться. Однако *каждая сателлитная ДНК возникла в результате амплификации очень короткой последовательности.* Эта последовательность либо представляет собой вариант предсуществовавшей сателлитной ДНК (как в случае с *D. virilis*), либо имеет иное происхождение.

Сателлитные ДНК постоянно возникают и исчезают из генома. Поэтому выявление эволюционных взаимосвязей сателлитов становится довольно сложной задачей. Существующая в настоящее время сателлитная ДНК могла произойти от некого предкового сателлита, который впоследствии был утерян. Важной особенностью сателлитов является то, что *они представляют собой протяженные участки ДНК, состоящие из коротких повторов, в пределах которых может сохраняться постоянство последовательности.*

Геном *D. virilis* содержит четыре родственных сателлита			
Сателлит	Превалирующая последовательность	Суммарная длина	Доля генома
I	ACAAACT TGTTTGA	$1{,}1 \times 10^7$	25%
II	ATAAACT TATTTGA	$3{,}6 \times 10^6$	8%
III	ACAAATT TGTTTAA	$3{,}6 \times 10^6$	8%
Криптический	AATATAG TTATATC		

РИС. 6.23 Все сателлитные ДНК *Drosophila virilis* родственны друг другу. Более 95% каждого сателлита состоит из тандемных повторов превалирующей последовательности

Еще одно отличительное свойство сателлитов состоит в ярко выраженной асимметрии нуклеотидных пар двух цепей ДНК. В примере с сателлитами *D.virilis*, изображенными на рис. 6.22, в каждой из основных фракций сателлитной ДНК одна цепь существенно богаче нуклеотидами T и G. Это увеличивает ее плавучую плотность таким образом, что во время денатурации ее **тяжелая цепь** (H) может быть отделена от комплементарной ей **легкой цепи** (L). Это свойство может быть полезным при определении нуклеотидной последовательности сателлитной ДНК.

6.13 Сателлитная ДНК млекопитающих состоит из иерархических повторов

Ключевые понятия:

- Сателлитная ДНК мышей эволюционировала путем дупликации и мутирования исходной короткой повторяющейся единицы в основную повторяющуюся единицу размером 234 пн, в половине, четверти и восьмой части которой можно распознать исходный повтор.

На примере грызунов показано, что у млекопитающих входящие в состав сателлитной ДНК последовательности демонстрируют сильные различия в структуре тандемного повтора. Доминирующий сателлитный повтор может быть обнаружен за счет его преобладания среди олигонуклеотидных фрагментов, получаемых при химической или ферментативной обработке ДНК. Заметим, что даже доминирующий повтор обычно составляет только небольшую часть от всех сателлитных фрагментов. Другие короткие повторы связаны с доминирующим повтором многочисленными нуклеотидными заменами, делециями и инсерциями.

Набор таких «родственных» вариантов короткого повтора может формировать более длинную повторяющуюся единицу, которая сама по себе тандемно повторяется с некоторыми изменениями. Таким образом, сателлитная ДНК млекопитающих состоит из иерархически организованных повторяющихся единиц. Эти более длинные повторяющиеся единицы формируют последовательности, эффективно ренатурирующие при проведении анализа методом реассоциации. Кроме того, их можно обнаружить при расщеплении ДНК рестриктазами.

Когда сателлитную ДНК расщепляют ферментом, сайт узнавания которого расположен в повторяющейся единице ДНК, для каждой из содержащих этот сайт

повторяющихся единиц обнаруживается сходный фрагмент. В действительности, при обработке геномной ДНК эукариот рестриктазами, основная ее часть вследствие случайного распределения сайтов рестрикции при электрофорезе в геле формирует размытое пятно (шмер). В свою очередь, сателлитная ДНК формирует (на фоне геномного шмера) четкие полосы, поскольку при расщеплении ДНК по расположенным через регулярные интервалы сайтам рестрикции образуется множество одинаковых фрагментов.

Прямое определение последовательности сателлитной ДНК может вызывать затруднения. Можно попытаться определить ее последовательность (секвенировать), используя отдельные полосы, образовавшиеся в результате обработки рестриктазами. К сожалению, при наличии существенных различий между отдельными повторяющимися единицами в одной и той же позиции разных повторов будут расположены разные нуклеотиды, поэтому сиквенсная электрофореграмма даст нечеткую картину. Если же отличия повторяющихся блоков незначительны — например, в пределах 2%, — то становится возможным определение усредненного нуклеотидного состава повторяющейся единицы.

Еще один подход заключается в клонировании отдельных участков сателлитов при помощи плазмидных векторов с целью последующего секвенирования клона. Сложность здесь состоит в том, что в результате рекомбинации в бактерии-хозяине последовательности сателлитной ДНК часто исчезают из химерных плазмид. Однако если клонирование проходит успешно, последовательность клонированного фрагмента можно определить довольно однозначно. Несмотря на то что таким образом можно определить реальную последовательность повторяющейся единицы или единиц, для выявления характера различий, типичных для сателлитной ДНК в целом, мы должны располагать данными о значительном числе таких индивидуальных последовательностей.

Любой метод прямого определения последовательности не дает исчерпывающей информации, поскольку все они ограничены участком, последовательность которого может быть определена данным методом за один раз. Повторение различающихся тандемных копий делает невозможным восстановление более длинной последовательности путем компьютерного сопоставления отдельных перекрывающихся рестриктных фрагментов.

Сателлитная ДНК мыши *M. musculus* расщепляется ферментом EcoRII на ряд фрагментов, включая основной мономерный фрагмент длиной 234 пн. Этот фрагмент, по-видимому, повторяется с небольшими вариациями на протяжении 60–70% от всей сателлитной ДНК мыши (расщепляемой на мономерные фрагменты данной рестриктазой). Проанализируем эту последовательность с точки зрения составляющих ее повторов меньшего размера.

На **РИС. 6.24** показаны два полуповтора этой последовательности. Изобразив последовательность длиной в 234 пн таким образом, чтобы первые 117 пн располагались над вторыми 117 пн, мы увидим, что эти две половины довольно похожи друг на друга. Они отличаются лишь по 22 позициям, что соответствует степени дивергенции в 19%. Это означает, что существующая в настоящее время повторяющаяся единица длиной 234 пн, по-видимому, возникла в прошлом в результате дупликации повторяющейся единицы длиной 117 пн, после чего дубликаты приобрели различия.

В пределах последовательности длиной 117 пн можно обнаружить две другие субъединицы. Каждая из них представляет собой четверть повтора (от сателлита в целом). Четыре четверти повтора выровнены друг относительно друга на **РИС. 6.25**. Две верхние последовательности представляют собой первый полуповтор последовательностей, изображенных на рис. 6.24, а две нижние последовательности составляют второй полуповтор. Видно, что дивергенция между четырьмя четвертями повтора увеличивается до 23 из 58 нуклеотидов, или до 40%. Первые три четверти повтора связаны между собой несколько лучше, а основной вклад в дивергенцию вносят изменения последней четверти повтора.

Рассматривая последовательности четвертьповторов, можно обнаружить, что каждый из них состоит из двух взаимосвязанных единиц (восьмых частей повтора), обозначенных на **РИС. 6.26** как α- и β-последовательности. Во всех α-последовательностях есть вставка одного нуклеотида (C), а во всех β-последовательностях — вставка тринуклеотида, сходного с общей консенсусной последовательностью. Это позволяет предположить, что четверти повтора произошли путем дупликации последовательности, похожей на консенсусную, после чего каждый из них приобрел изменения и сформировал компоненты, обозначенные нами как α и β. В дальнейшем между тандемно повторяющимися последователь-

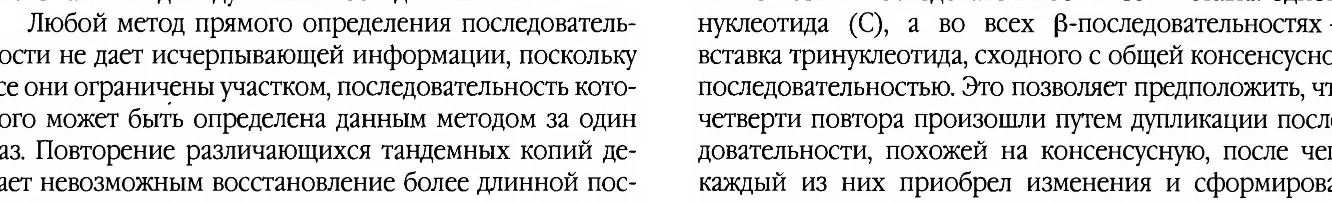


РИС. 6.24 Повторяющаяся единица сателлитной ДНК мыши длиной 234 пн состоит из двух почти идентичных половин (на рисунке эти половины расположены друг под другом, демонстрируя, что в подавляющем большинстве положений их нуклеотиды совпадают)



РИС. 6.25 Сопоставление четвертей повтора показывает большое сходство между первой и второй половинами каждого полуповтора. Положения, занимаемые во всех четырех четвертях одними и теми же нуклеотидами, отмечены голубым шрифтом и темно-серым фоном

ностями αβ накапливались изменения, что и привело к образованию отдельных четверть- и полуповторов, существующих в настоящее время. Дивергенция между восьмыми частями повтора составляет 19/31=61%.

Прямой анализ консенсусной последовательности приведен на РИС. 6.27, где показано, что нынешняя последовательность сателлитной ДНК может быть представлена в виде производных от последовательности длиной 9 пн. Как отмечено в нижней части рисунка, мы можем распознать три варианта такой последовательности в составе сателлита. Если в одном из повторов наиболее часто встречающееся основание заменить в двух положениях на следующее по частоте встречаемости, то мы получим три близкие последовательности длиной 9 пн:

GAAAAACGT
GAAAAATGA
GAAAAAACT

Происхождение сателлита легко объясняется процессом амплификации одного из этих трех наномеров. Общая консенсусная последовательность нынешнего сателлита может быть записана как GAAAAA$^{AG}_{TC}$T, и она фак-



РИС. 6.26 Сопоставление восьмых частей повтора показывает, что каждая четверть повтора состоит из похожих друг на друга половин, α и β. Консенсус одной восьмой части 234 пн повтора строится стандартным способом (в каждое положение ставится наиболее часто встречающийся нуклеотид). Консенсус может быть эквивалентом «предковой» последовательности, от которой дивергировали α и β

Консенсус сателлитов мыши имеет длину 9 пн

```
          G G A C C T
G G A A T A T G G C
G A G A A A A C T
G A A A A T C A C
G G A A A A T G A
G A A A T C A C T
T T A G G A C G T
G A A A T A T G G C
G A G A^G A A A C T
G A A A A A G G T
G G A A A A T^T T A
G A A A T* C A C T
G T A G G A C G T
G G A A T A T G G C
A A G A A A A C T
G A A A A T C A T
G G A A A A T G A
G A A A C* C A C T
T G A C G A C T T
G A A A A A T G A C
G A A A T C A C T
A A A A A A C G T
G A A A A A T G A
G A A A T* C A C T
G A A
```

G_{20} A_{16} A_{21} A_{20} A_{12} A_{17} T_8 G_{11} A_5

T_7 C_5 A_8 C_9 T_{15}

C_7

* Вставка триплета в β-последовательность

C в позиции 10 является дополнительным нуклеотидом α-последовательности

РИС. 6.27 Написание последовательности сателлита, как бы состоящей из 9 пн повторов, помогает подтвердить наличие у всех этих «повторов» реального консенсуса

тически представляет собой смесь из трех приведенных выше повторов длиной 9 пн.

Свойства сателлитной ДНК определяются строением усредненной последовательности ее мономерного фрагмента. При расщеплении рестриктазой обнаруживается самая длинная повторяющаяся единица сателлита размером 234 пн. Единица реассоциации отдельных цепей денатурированной сателлитной ДНК — это, по-видимому, полуповтор размером 117 пн, поскольку фрагменты размером 234 пн могут реассоциировать как полностью, так и наполовину (в последнем случае первый полуповтор одной цепи ренатурирует со вторым полуповтором другой).

До сих пор мы рассматривали сателлитную ДНК как повтор идентичных фрагментов размером 234 пн. Хотя эта единица и составляет большую часть сателлита, встречаются также и ее варианты. Некоторые из них распределены в сателлитной ДНК случайно, в то время как другие организованы в виде кластеров.

Мы подразумевали существование этих вариантов, когда описывали исходный материал для анализа последовательности как мономерный фрагмент. При расщеплении сателлитной ДНК ферментом, для которого в последовательности размером 234 пн есть только один сайт рестрикции, образуются также димеры, тримеры и тетрамеры этого фрагмента. Они возникают, когда повторяющаяся единица утрачивает сайт рестрикции в результате мутации.

Мономерная единица размером 234 пн формируется в том случае, если каждый из двух прилегающих повторов содержит сайт рестрикции. Димер формируется, когда одна единица утратила сайт, а тример — когда обе прилегающие единицы утратили сайты рестрикции. С помощью некоторых рестриктаз большая часть сателлитной ДНК расщепляется на повторяющиеся серии фрагментов, как, например, на РИС. 6.28. Уменьшение числа димеров, тримеров и т. д. свидетельствует о случайном распределении повторов, утерявших сайт распознавания в результате мутации.

Другие рестриктазы ведут себя иначе по отношению к сателлитной ДНК. Они продолжают формировать одинаковые серии полос. Однако они расщепляют лишь небольшую часть сателлитной ДНК, скажем 5–10%. Это значит, что некий участок сателлита содержит множество повторяющихся единиц с определенным сайтом распознавания. По-видимому, серии повторов в этом участке ДНК произошли от единственного варианта повтора, содержащего этот сайт распознавания (хотя обычно су-



РИС. 6.28 Расщепление сателлитной ДНК мыши ферментом EcoRII дает серию интенсивных полос (1, 2, 3), которые представляют собой мультимеры 234 пн повторов. Кроме них просматриваются также минорные серии полос (½, 1½, 2½), кратных половинам повторов (см. текст на этой странице). Крайняя слева полоса соответствует фракции сателлитной ДНК, устойчивой к расщеплению

ществует несколько членов кластера, утерявших сайт в результате мутации).

Сателлитная ДНК подвержена воздействию неравной рекомбинации, в особенности когда в повторяющейся единице есть внутренние повторы. Вернемся к нашему кластеру, состоящему из повторов ab. Предположим, что компоненты повторяющейся единицы a и b в значительной степени схожи. В таком случае два кластера можно выровнять таким образом, чтобы последовательности типа a одного кластера соответствовали последовательностям типа b другого. Частота этого события будет зависеть от степени сходства между двумя вариантами повторяющейся единицы. Описанным образом происходит реассоциация цепей денатурированной сателлитной ДНК мыши *in vitro*.

В результате рекомбинации со сдвигом происходит увеличение числа повторяющихся единиц, вовлеченных в реакцию:

xabababababababababababababababy
x
xababababababababababababababababy
↓
xababababababababaababababababababy
+
xabababababababababbabababababababy

В верхнем рекомбинантном кластере единица «ab» была заменена на единицу «aab», а в нижнем — на единицу «b».

Такая рекомбинация объясняет рестрикционную картину, получающуюся при обработке мышиной сателлитной ДНК рестриктазами. На рис. 6.28 показано, что, помимо серий целых повторяющихся единиц, в меньшем количестве представлены фрагменты длиной в $^1/_2$, $1^1/_2$, $2^1/_2$ и $3^1/_2$ от длины повторяющейся единицы. Предположим, что в предыдущем примере ab представляет повторяющуюся единицу мышиной сателлитной ДНК размером 234 пн, образовавшуюся при расщеплении по сайту, расположенному в сегменте b. Фрагменты a и b соответствуют полуповторам длиной 117 пн.

Тогда в верхнем рекомбинантном кластере единица «aab» сформирует фрагмент, в полтора раза больше размера обычной повторяющейся единицы. В нижнем рекомбинантном кластере единица «b» формирует фрагмент по размеру в два раза меньше обычной повторяющейся единицы. (Короткие фрагменты, состоящие из полуповторов, формируются так же, как и более длинные фрагменты, состоящие из целых повторов. Это происходит из-за утраты некоторыми повторяющимися единицами сайтов рестрикции в результате мутации.)

И наоборот, наличие серий полуповторов на геле указывает на то, что повторяющаяся единица длиной 234 пн состоит из двух полуповторов, достаточно сходных для того, чтобы иногда принимать участие в рекомбинации. На рис. 6.28 можно заметить несколько слабовыраженных пиков, соответствующих фрагментам, размер которых составляет $^1/_4$ и $^3/_4$ от размера основной повторяющейся единицы. Такие фрагменты образуются так же, как и фрагменты, составленные из полуповторов, когда рекомбинация возникает между кластерами, выровненными со сдвигом на четверть длины. Меньшая степень сходства между четвертями повтора по сравнению с полуповторами объясняет уменьшение частоты пиков, соответствующих фрагментам размером $^1/_4$ и $^3/_4$ по сравнению с пиками полуповторов.

6.14 Мини-сателлиты можно использовать при генетическом картировании

Ключевые понятия:

- Вариабельность длины микросателлитов и мини-сателлитов отдельных геномов можно использовать для определения степени их биологического родства. Для этого достаточно показать, что 50% исследованных микросателлитов и мини-сателлитов индивидуума получены от определенного родителя.

В геномах млекопитающих часто встречаются последовательности, которые по строению напоминают сателлиты (состоят из тандемных повторов), но при этом значительно короче — длина их может составлять от 5 до 50 повторов. Их обнаружили случайно в виде очень изменчивых по размеру фрагментов при изучении библиотек геномной ДНК человека. Изменчивость становится заметной, когда популяция содержит фрагменты, представляющие один и тот же регион генома, но различающиеся по размеру; при индивидуальном тестировании индивидуумов выясняется, что между генами существует значительный полиморфизм и что в популяции присутствует множество аллелей таких сателлитов.

Термин **микросателлит** обычно используют, когда длина повторяющейся единицы составляет менее 10 пн, а мини-сателлит — когда длина единицы находится в диапазоне от 10 до 100 пн (хотя точное определение отсутствует). В общем виде такие последовательности называют **варьирующими по числу тандемными повторами** (variable number tandem repeats, VNTR).

Существование вариаций микросателлитов или мини-сателлитов между индивидуальными геномами связано с тем, что различные аллели содержат разное число повторяющихся единиц. Например, мини-сателлит с длиной повтора в 64 пн распределен в популяции следующим образом:

7%	18 повторов
11%	16 повторов
43%	14 повторов
36%	13 повторов
4%	10 повторов

Скорость генетических изменений в последовательностях мини-сателлитов довольно велика, составляя око-

ло 10^{-4} на тысячу пар нуклеотидов. (Считается, что частота замен в данном локусе пропорциональна его длине.) Эта скорость примерно в 10 раз выше скорости гомологичной рекомбинации при мейозе (то есть скорости рекомбинации любой случайно взятой последовательности ДНК).

Высокая изменчивость мини-сателлитов делает их удобным маркером для генетического картирования, поскольку существует высокая вероятность того, что по аллелям данного локуса индивиды будут отличаться. Пример картирования с помощью мини-сателлитов изображен на РИС. 6.29. На нем изображен крайний случай, когда два индивида гетерозиготны по мини-сателлитному локусу, и, по сути, все четыре аллеля различны. Все потомки получают по одному аллелю от каждого из родителей, что позволяет однозначно определить, от кого пришел каждый из аллелей потомка. В контексте генетики человека, изображенный на этом рисунке мейоз высокоинформативен в силу существования вариации между аллелями.

Одно семейство мини-сателлитов в геноме человека имеет общую исходную последовательность. Эта последовательность представляет собой G-C богатый участок длиной в 10–15 пн, демонстрирующий асимметрию в распределении пуриновых/пиримидиновых оснований на двух цепях. Каждый отдельный мини-сателлит представлен собственным вариантом повторяющейся последовательности, однако около 1000 мини-сателлитов можно детектировать методом саузерн-блот гибридизации с пробой, содержащей базовую последовательность.

Представьте изображенный на рис. 6.29 пример, только многократно расширенный за счет существования множества подобных последовательностей. Результатом существования множества локусов со множеством вариантов в каждом отдельном локусе становится существование уникального паттерна для каждого индивида. Это делает возможным однозначное определение родства между родителями и потомством за счет того, что каждый ребенок получает ровно 50% локусов от каждого из родителей. На этом основан метод «генетических отпечатков пальцев».

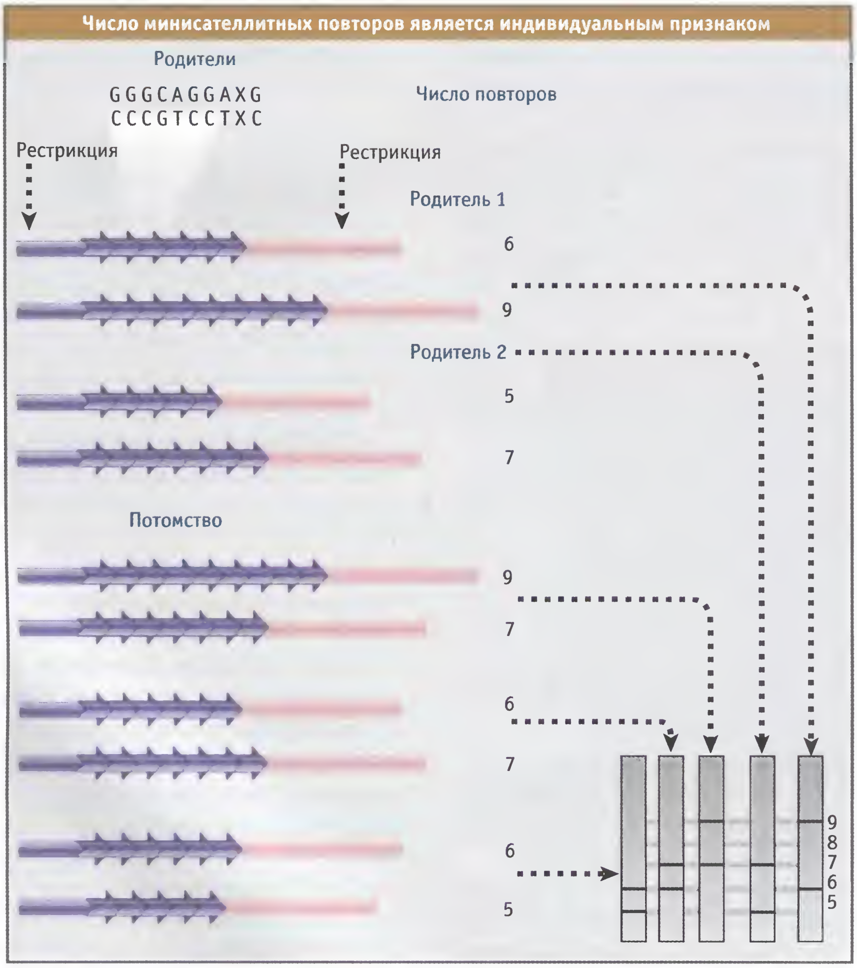


РИС. 6.29 Аллели локуса-мини-сателлита могут различаться по числу повторов и, благодаря этому, давать рестрикционные фрагменты, различающиеся по длине. Используя мини-сателлит, аллели которого у двух родителей отличаются, можно исследовать паттерн наследования

И микросателлиты, и мини-сателлиты нестабильны, хотя и по разным причинам. Микросателлиты накапливают неспаренные основания в результате проскальзывания во время репликации, приводящего к увеличению числа повторов (РИС. 6.30). В исправлении таких изменений важную роль играют репарационные системы, распознающие неспаренные основания. Мутации в генах системы репарации играют значительную роль в развитии рака, поскольку опухолевые клетки часто демонстрируют изменчивость в последовательностях микросателлитов. Мини-сателлиты, в свою очередь, претерпевают неравный кроссинговер между повторами (по механизму, очень схожему с тем, что мы обсудили для сателлитной ДНК (рис. 6.3)). Высокая изменчивость мини-сателлитов делает их «горячей точкой» рекомбинации при мейозе. Рекомбинация в данном случае не является рекомбинацией между фланкирующими последовательностями, она принимает сложную форму, когда новый мутантный аллель одновременно получает часть сестринской хроматиды и часть другой (гомологичной) хромосомы.

Остается неясным, при какой длине повторяющейся единицы происходит переход от механизма создания новых вариантов путем сдвига репликации к рекомбинационному механизму.

6.15 Резюме

Почти все гены входят в состав того или иного семейства, объединяемого наличием схожих последовательностей экзонов между членами семейства. Эволюция семейств обусловлена дупликацией генов, с последующим накоплением отличий между копиями. Некоторые копии приобретают инактивирующие мутации и превращаются в псевдогены, не выполняющие никакой функции. Кроме того, псевдогены могут возникать в виде ДНК-копий молекул мРНК.

Эволюционирующие гены могут составлять единый кластер, но могут быть и распределены по разным

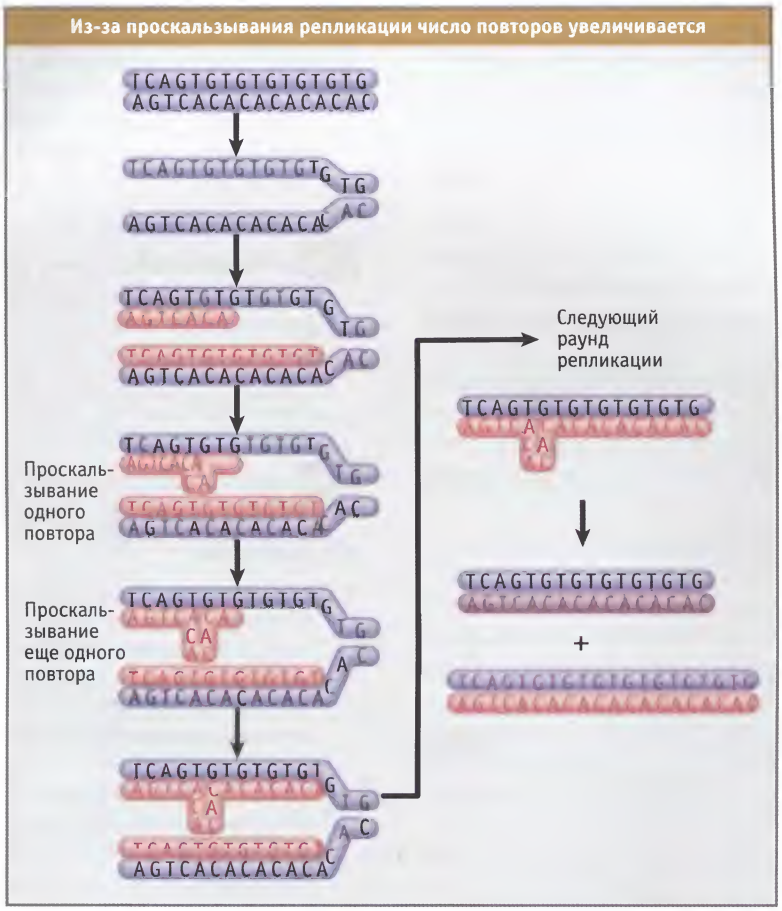


РИС. 6.30 Под проскальзыванием репликации подразумевают сдвиг дочерней цепи вдоль матрицы на один повтор назад. Каждое такое событие добавляет дочерней цепи один лишний повтор (его нет в матричной цепи, так что в дочернем дуплексе лишний повтор будет иметь форму одноцепочечной петли). В следующем раунде репликации дочерняя цепь с лишним повтором послужит матрицей и даст начало полноценному дуплексу ДНК с лишним повтором

местам хромосомы (за счет рекомбинации). Структура существующих кластеров иногда позволяет сделать заключение о событиях, приведших к их формированию. Эти события влияют скорее на последовательность, чем на функцию и, таким образом, затрагивают как активные гены, так и псевдогены. Мутации возникают чаще в молчащих сайтах, чем в сайтах замен (где они вызывают аминокислотные замены). Уровень дивергенции в сайтах замен можно использовать для создания модели эволюционных часов, которые можно откалибровать по проценту дивергенции за миллион лет. Эти часы позволяют оценить время разделения двух любых членов генного семейства.

Тандемные кластеры состоят из множества копий повторяющихся единиц, включающих транскрибируемую последовательность (последовательности) и нетраскрибируемый спейсер (спейсеры). Кластеры генов рРНК кодируют одну молекулу-предшественник рРНК. Сохранение активных генов в кластерах зависит от таких механизмов, как генная конверсия и неравный кроссинговер, приводящих к быстрому распространению мутации на весь кластер и «подставлению» генов под действие отбора.

Сателлитная ДНК состоит из коротких тандемно повторяющихся последовательностей. Ее индивидуальные физические свойства объясняются особенностями нуклеотидного состава. Сателлитная ДНК сконцентрирована в гетерохроматине центромер, однако ее функция (если таковая есть) неизвестна. Отдельные повторяющиеся единицы сателлитной ДНК членистоногих идентичны. Сателлиты млекопитающих похожи между собой и могут быть организованы иерархически, что свидетельствует о вкладе в их эволюцию таких процессов, как амплификация и дивергенция (случайно выбранных повторов).

Неравный кроссинговер является основным фактором, влияющим на структуру сателлитной ДНК. Кроссоверная фиксация объясняет возможность распространения конкретного варианта повторяющегося элемента на весь кластер.

Мини-сателлиты и микросателлиты состоят из более коротких (чем сателлиты) повторяющихся последовательностей: микросателлиты — менее 10 пн, а мини-сателлиты — от 10 до 100 пн. Число повторяющихся единиц в кластере обычно составляет от 5 до 50. Между отдельными геномами внутри вида существует значительная изменчивость по числу повторов в кластере. Число повторов микросателлитов в кластере может меняться в результате сдвига во время репликации; а частота сдвига зависит от работы систем, распознающих и исправляющих ошибки при репликации ДНК. Число повторов мини-сателлитов варьирует в результате событий, напоминающих неравную рекомбинацию. Изменчивость числа повторов может быть использована для определения родственных связей при проведении анализа, известного как метод «генетических отпечатков пальцев».

Литература

6.2 Дупликация генов — основная движущая сила эволюции

Исследование

Bailey, J. A., Gu, Z., Clark, R. A., Reinert, K., Samonte, R. V., Schwartz, S., Adams, M. D., Myers, E. W., Li, P. W., and Eichler, E. E. (2002). Recent segmental duplications in the human genome. *Science* 297, 1003–1007.

6.3 Кластеры генов глобинов возникли путем дупликации и дивергенции

Обзор

Hardison, R. (1998). Hemoglobins from bacteria to man: evolution of different patterns of gene expression. *J. Exp. Biol.* 201, 1099–1117.

6.5 Дивергенция повторяющихся последовательностей отражает скорость накопления нейтральных замен

Исследование

Waterston, R. H. et al. (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature* 420, 520–562.

6.6 Псевдогены — тупики эволюции

Исследование

Hirotsune, S., Yoshida, N., Chen, A., Garrett, L., Sugiyama, F., Takahashi, S., Yagami, K., Wynshaw-Boris, A., and Yoshiki, A. (2003). An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. *Nature* 423, 91–96.

6.10 Рекомбинационная фиксация может поддерживать идентичность повторов

Обзор

Charlesworth, B., Sniegowski, P., and Stephan, W. (1994). The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371, 215–220.

6.14 Мини-сателлиты можно использовать при генетическом картировании

Исследование

Jeffreys, A. J., Murray, J., and Neumann, R. (1998). High-resolution mapping of crossovers in human sperm defines a minisatellite-associated recombination hotspot. *Mol. Cell* 2, 267–273.

Jeffreys, A. J., Royle, N. J., Wilson, V., and Wong, Z. (1988). Spontaneous mutation rates to new length alleles at tandem-repetitive hypervariable loci in human DNA. *Nature* 332, 278–281.

Jeffreys, A. J., Tamaki, K., MacLeod, A., Monckton, D. G., Neil, D. L., and Armour, J. A. (1994). Complex gene conversion events in germline mutation at human minisatellites. *Nat. Genet.* 6, 136–145.

Jeffreys, A. J., Wilson, V., and Thein, S. L. (1985). Hypervariable minisatellite regions in human DNA. *Nature* 314, 67–73.

Strand, M., Prolla, T. A., Liskay, R. M., and Petes, T. D. (1993). Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature* 365, 274–276.

Матричная РНК

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

- Один из путей деградации мРНК дрожжей основан на ее расщеплении экзонуклеазами рестрикции в направлении от 5′- к 3′-концу.
- Другой путь деградации основан на действии комплекса из нескольких экзонуклеаз, работающих в направлении от 3′- к 5′-концу.
- Деаденилаза клеток животных может напрямую связываться с 5′-кэпом.

7.14 Нонсенс-мутации запускают «систему надзора»

- Нонсенс-мутации являются причиной деградации мРНК.
- Гены, кодирующие систему деградации, были найдены у дрожжей и червей.

7.15 РНК эукариот транспортируется через мембрану

- Трансмембранный транспорт РНК осуществляется в составе рибонуклеопротеина.
- Все РНК эукариот, работающие в цитоплазме, должны быть экспортированы из ядра.
- ТРНК и РНК-компонент рибонуклеазы импортируются в митохондрию.
- мРНК растений может осуществлять длинные путешествия между клетками.

7.16 Местонахождение мРНК может быть четко детерминировано

- мРНК Ash1 дрожжей формирует рибонуклеопротеин, связывающийся с миозиновым мотором.
- Мотор транспортирует ее вдоль актиновых филаментов в дочернюю клетку.
- В дочерней клетке она заякоривается и транслируется. Таким образом, белок синтезируется только в дочерней клетке.

7.17 Резюме

7.1 Введение

РНК является основным участником процесса экспрессии генов. Сначала ее характеризовали как промежуточный продукт синтеза белка, однако с открытием других видов РНК стало понятно, что они играют структурные и функциональные роли на других этапах экспрессии гена. Участие РНК во множестве процессов, так или иначе связанных с экспрессией гена, поддерживает распространенную точку зрения о том, что именно РНК была исходным компонентом, осуществлявшим сохранение и экспрессию генетической информации на заре эволюции. Многие функции, первоначально принадлежавшие РНК, были впоследствии перенесены на белки, изменчивость которых могла привести к увеличению эффективности выполняемой функции.

На РИС. 7.1 изображены три основных класса РНК, напрямую вовлеченных в процесс синтеза белка:

- Матричная (информационная) РНК (мРНК) является промежуточной копией последовательности ДНК, несущей информацию о строении белка.
- Транспортные РНК **(тРНК)** — это сравнительно небольшие молекулы, доставляющие соответствующие аминокислоты к кодонам мРНК.
- Рибосомальные РНК **(рРНК)** — компоненты рибосом — крупных рибонуклеопротеиновых комплексов, содержащих множество белков и РНК и осуществляющих синтез полипептидной цепи из аминокислот.

В каждом из этих случаев РНК играет отдельную роль. Важной особенностью матричной РНК является ее последовательность: каждый триплет нуклеотидов в пределах кодирующей части мРНК представляет аминокислоту в соответствующем белке. Однако структура мРНК и последовательности, расположенные по концам кодирующего региона, могут играть важную роль в регуляции ее активности, а следовательно, и в регуляции количества производимого ею белка.

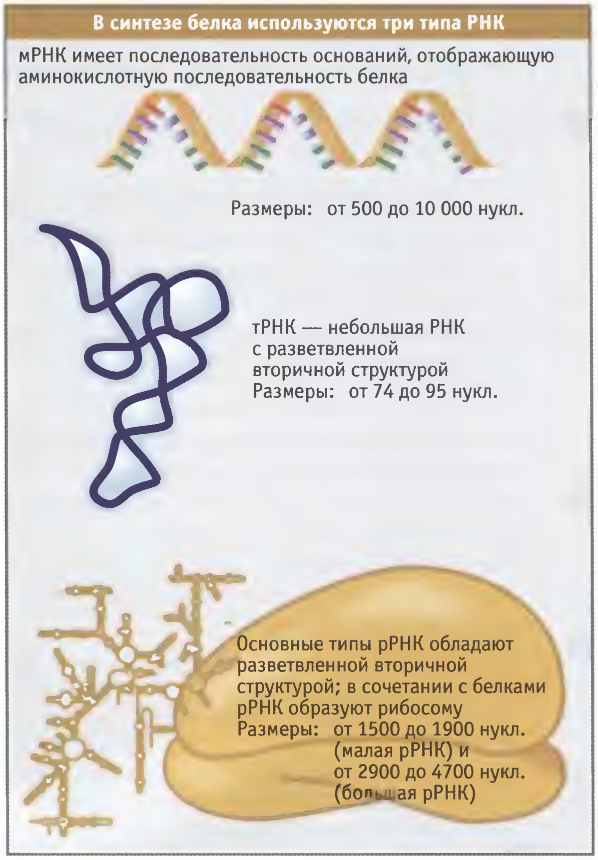


РИС. 7.1 Для экспрессии любого гена требуются РНК трех типов: мРНК (носитель кодирующей последовательности), тРНК (поставщик аминокислот в строгом соответствии с кодонами), и рРНК (главный компонент рибосомы, создающий условия для синтеза белка)

Функции тРНК также определяются ее структурой и последовательностью: в тРНК важна как пространственная структура, так и ее способность к попарному взаимодействию с нуклеотидами другой РНК (мРНК). На первом этапе пространственная структура тРНК распознается ферментом, который определяет сов-

местимость РНК с конкретной аминокислотой. Если они совместимы, между ними формируется связь, приводящая к образованию аминоацил-тРНК, которая впоследствии и используется при синтезе белка. Специфичность аминоацил-тРНК определяется принципом комплементарности нуклеотидов, составляющих антикодон, и соответствующего триплета нуклеотидов мРНК, кодирующего данную аминокислоту.

В случае с рРНК мы наблюдаем другой тип активности. Одна из ее функций — структурная рРНК формирует каркас, к которому крепятся рибосомальные белки. Помимо этого, рРНК принимает непосредственное участие в функционировании рибосомы. Одной из основных функций рибосомы является ее способность катализировать формирование пептидной связи, посредством которой аминокислота прикрепляется к белку. Эта функция реализуется одной из рРНК.

Важно понимать, что, поскольку мы признаем значимость роли РНК в процессе синтеза белка, мы должны рассматривать ее как активного участника этого процесса, который, помимо прочего, может быть мишенью для регулирующего воздействия белков и других РНК. Напомним, что РНК могли быть основой первичного аппарата синтеза белка. Важным фактором любой активности РНК является принцип спаривания оснований, играющий ключевую роль как в формировании ее вторичной структуры, так и в специфичности ее взаимодействия с другими молекулами РНК. Кодирующая функция мРНК исключительно важна, однако тРНК и рРНК являются представителями гораздо более широкого класса некодирующих РНК, выполняющих разнообразные функции в процессе экспрессии гена.

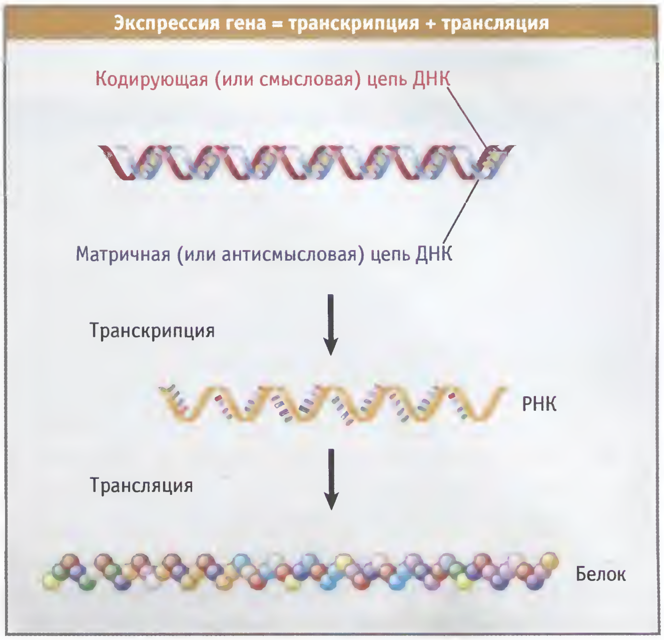


РИС. 7.2 При транскрипции образуется РНК, комплементарная матричной цепи ДНК и (по последовательности оснований) идентичная кодирующей цепи ДНК. При трансляции каждый триплет азотистых оснований этой РНК переводится в одну аминокислоту синтезируемого белка. Три полных витка двойной спирали ДНК содержат тридцать пн, которые, будучи кодирующими, соответствуют десяти аминокислотам

7.2 мРНК, полученная в результате транскрипции, транслируется в аминокислотную последовательность белка

Ключевые понятия:

- Матрицей для синтеза РНК служит лишь одна из двух цепей ДНК.

Экспрессия гена осуществляется в два этапа.

- В результате транскрипции формируется одноцепочечная РНК, идентичная последовательности одной из цепей ДНК-дуплекса.
- В результате трансляции нуклеотидная последовательность мРНК переводится в последовательность аминокислот, составляющих белок. Последовательность мРНК транслируется не целиком, однако каждая мРНК содержит по крайней мере один кодирующий участок, связанный с последовательностью белка генетическим кодом: каждый триплет нуклеотидов (кодон) кодирует одну аминокислоту.

Лишь одна из цепей ДНК-дуплекса транскрибируется в мРНК. Поэтому в ДНК-дуплексе принято различать антисмысловую и смысловую (кодирующую) цепи (РИС. 7.2).

- Цепь ДНК, регулирующая синтез мРНК посредством комплементарного спаривания нуклеотидов, называют матричной или **антисмысловой** цепью. (Термин *антисмысловая* используется для описания последовательности ДНК или РНК, комплементарной мРНК.)
- Другая цепь ДНК содержит последовательность, *идентичную* последовательности мРНК (за исключением того, что на месте U в ДНК расположен T), и ее называют **кодирующей,** или **смысловой** цепью.

В этой главе мы обсудим мРНК и ее использование в качестве матрицы для синтеза белков. В гл. 8 «Синтез белка» мы обсудим сам процесс синтеза белка. В гл. 9 «Расшифровка генетического кода» мы опишем, как с помощью генетического кода интерпретируется значение последовательности мРНК. В гл. 10 «Локализация белка» мы рассмотрим вопрос, касающийся механизмов правильного расположения вновь синтезированных белков в клетке.

7.3 Транспортная РНК по форме напоминает лист клевера

Ключевые понятия:

- тРНК состоит из последовательности 74–95 нуклеотидов, формирующих вторичную структуру с четырьмя плечами, напоминающими лист клевера (у более длинных тРНК может формироваться дополнительная петля).
- При возникновении **эфирной** связи между 2′- или 3′-OH-группой рибозы концевого аденина акцепторной петли тРНК и COOH-группой аминокислоты формируется аминоацил-тРНК.
- За специфичность аминоацил-тРНК ответственна только последовательность антикодона.

мРНК можно отделить от аппарата, ответственного за ее трансляцию, используя бесклеточные системы синтеза белков *in vitro*. *Система синтеза белка одной клетки может транслировать мРНК клетки другого типа, а значит, и генетический код, и аппарат трансляции универсальны.*

Каждый триплет нуклеотидов в мРНК кодирует аминокислоту. Несоответствие структуры тринуклеотида и аминокислоты поднимает вопрос о способе взаимодействия кодонов и аминокислот. Адаптором (переходником), необходимым для этого взаимодействия, является транспортная РНК (тРНК). тРНК обладает двумя важными свойствами:

- С ней может *ковалентно* связываться одна аминокислота.
- Она несет последовательность из трех нуклеотидов **(антикодон)**, которая *комплементарна кодону, представляющему данную аминокислоту*. Антикодон позволяет тРНК распознавать кодон посредством комплементарных взаимодействий нуклеотидов.

У всех тРНК одинаковая вторичная и третичная структуры. Вторичную структуру тРНК можно представить в виде **листа клевера**, где комплементарные участки формируют **шпильки**, а некомплементарные — одноцепочечные **петли** (РИС. 7.3). Шпильки с петлями носят название **плечей** тРНК. Их последовательности содержат нестандартные основания, сформированные в результате модификаций четырех обычных нуклеотидов после синтеза полинуклеотидной цепи.

Строение листа клевера более детально рассмотрено на РИС. 7.4. Четыре основных плеча получили название в соответствии со своей структурой и функцией:

- **Акцепторный конец** состоит из шпильки, заканчивающейся неспаренным участком, свободные 2′- или 3′-OH-группы которого могут связываться с аминокислотой.
- **TψC-плечо** получило свое название из-за наличия такого триплета (ψ означает модифицированное основание — псевдоуридин).
- **Антикодоновое плечо** (антикодон) в центре петли всегда содержит антикодон.
- **D-плечо** названо так из-за наличия в его составе дигидроуридина (другого модифицированного нуклеотида тРНК).

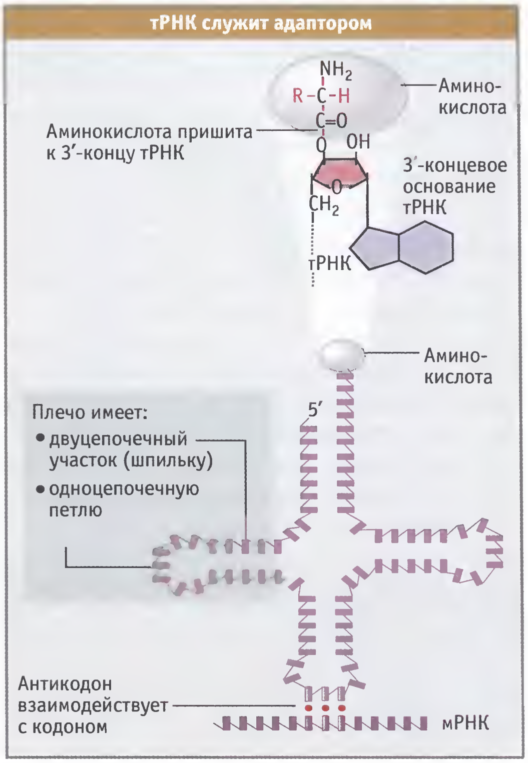


РИС. 7.3 Любая тРНК обладает свойствами адаптора с двойной специфичностью: к аминокислоте и к кодону мРНК. Аминокислота пришита ковалентной связью к 3′-концевому аденозину тРНК. Антикодон тРНК взаимодействует с кодоном мРНК по принципу комплементарности

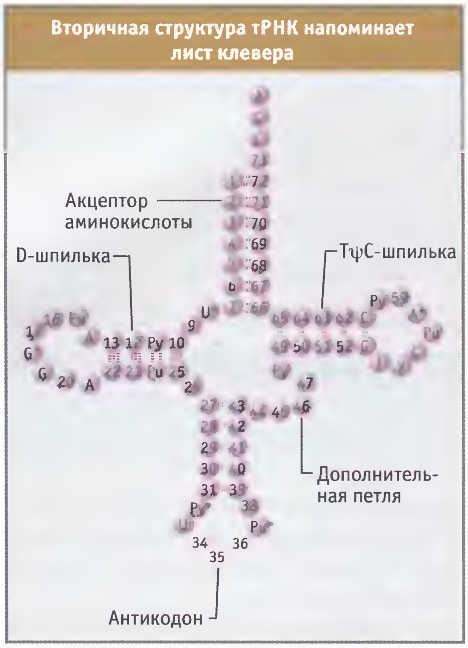


РИС. 7.4 В клеверном листе тРНК есть положения, занимаемые инвариантными и полуинвариантными рибонуклеотидами, а также консервативный набор комплементарных участков рибонуклеотидной последовательности

- **Дополнительная петля** расположена между TψC-рукой и рукой антикодона и по длине варьирует от 3 до 21 нуклеотида.

Система нумерации тРНК демонстрирует постоянство их структуры. Позиции нумеруются в направлении от 5′- к 3′-концу в соответствии с наиболее часто встречающейся структурой тРНК, содержащей 76 нуклеотидов. Вообще длины тРНК варьируют в диапазоне от 74 до 95 нуклеотидов. Существование тРНК разной длины объясняется изменчивостью D-плеча и дополнительного плеча.

Спаривание оснований, приводящее к сохранению вторичной структуры тРНК, изображено на рис. 7.4. В пределах данной тРНК большая часть пар представлена A-U и G-C, однако изредка встречаются пары G-U, G-ψ или A-ψ. Дополнительные типы пар нуклеотидов менее стабильны, нежели обычные пары, однако и они обладают потенциалом, достаточным для формирования двухцепочечных участков РНК.

При сравнении последовательностей тРНК видно, что некоторые позиции оснований **инвариантны** (или **консервативны**); почти всегда в консервативной позиции расположен определенный нуклеотид. Некоторые позиции описывают как **полуинвариантные** (или **полуконсервативные**), поскольку по ним могут находиться основания одного типа (пуриновые или пиримидиновые), однако это место может занимать любое основание данного типа.

Когда тРНК *загружена* аминокислотой, соответствующей ее антикодону, ее называют **аминоацил-тРНК**. Аминокислота крепится сложноэфирной связью своей карбоксильной группой к 2′- или 3′-гидроксильной группе рибозы 3′-концевого нуклеотида тРНК (это всегда аденин). Процесс загрузки тРНК катализируется ферментом **аминоацил-тРНК-синтазой**. Всего насчитывается (по крайней мере) 20 аминоацил-тРНК-синтаз. Каждая из них распознает одну аминокислоту и все тРНК, к которым эта аминокислота может быть присоединена.

Каждой аминокислоте соответствует как минимум одна тРНК (обычно больше). В названии тРНК верхним индексом указывают трехбуквенную аббревиатуру аминокислоты. Если одной и той же аминокислоте соответствует более чем одна тРНК, в нижнем индексе названия указывают цифры, позволяющие различать их. Например, две тРНК для тирозина будут называться $тРНК_1^{Tyr}$ и $тРНК_2^{Tyr}$. К названию тРНК, несущей аминокислоту (аминоацил-тРНК), добавляют префикс, обозначающий аминокислоту, то есть Ala-тРНК — это $тРНК^{Ala}$, несущая свою аминокислоту.

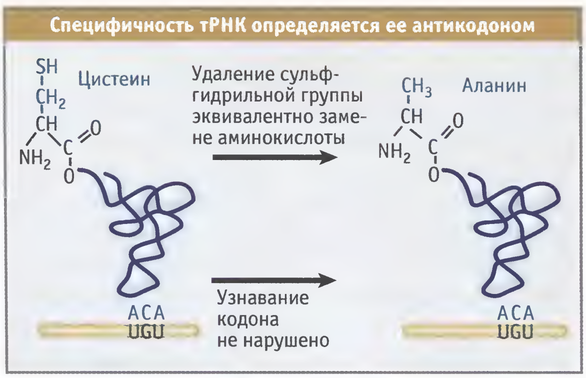


РИС. 7.5 Значение тРНК определяется антикодоном, но не аминокислотой

Достаточно ли одной последовательности антикодона для распознавания правильного кодона? Классический эксперимент, иллюстрирующий ответ на этот вопрос, изображен на РИС. 7.5. Восстановительное десульфурирование превращает цистеинил-$тРНК^{Cys}$ в аланил-$тРНК^{Cys}$. тРНК несет антикодон, соответствующий кодону UGU. Модификация аминокислоты не влияет на специфичность взаимодействия кодон-антикодон, то есть в белок вместо цистеина встроится аланин. *Аминокислота не играет роли в специфичности связывания с тРНК, она определяется исключительно антикодоном.*

7.4 Акцепторный конец и антикодон расположены по краям пространственной структуры

Ключевые понятия:

- Лист клевера формирует L-образную третичную структуру с акцепторной рукой на одном конце и рукой антикодона на другом.

Вторичная структура каждой тРНК формирует компактную L-образную третичную структуру, в которой 3′-конец, связывающийся с аминокислотой, удален от антикодона, взаимодействующего с мРНК. Все тРНК имеют одинаковую общую пространственную структуру, но, тем не менее, имеют индивидуальные отличия.

Спаренные двухцепочечные шпильки, характерные для вторичной структуры, сохраняются и в третичной структуре, однако их взаимное расположение в трехмерной системе координат формирует две двойные спирали, расположенные под прямым углом друг к другу, как показано на РИС. 7.6. Акцепторный участок и TψC-шпилька формируют одну длинную двойную спираль с разрывом. D-шпилька и антикодон формируют другую двойную спираль, также с разрывом. Участок между двойными спиралями, где L-образная структура совершает поворот, содержит TψC- и D-петли. Таким образом, аминокислота расположена на конце одного из плечей L-образной структуры, а петля антикодона — на конце другого.

Пространственная структура формируется за счет водородных связей между основаниями, которые не участвуют в формировании вторичной структуры. Эти

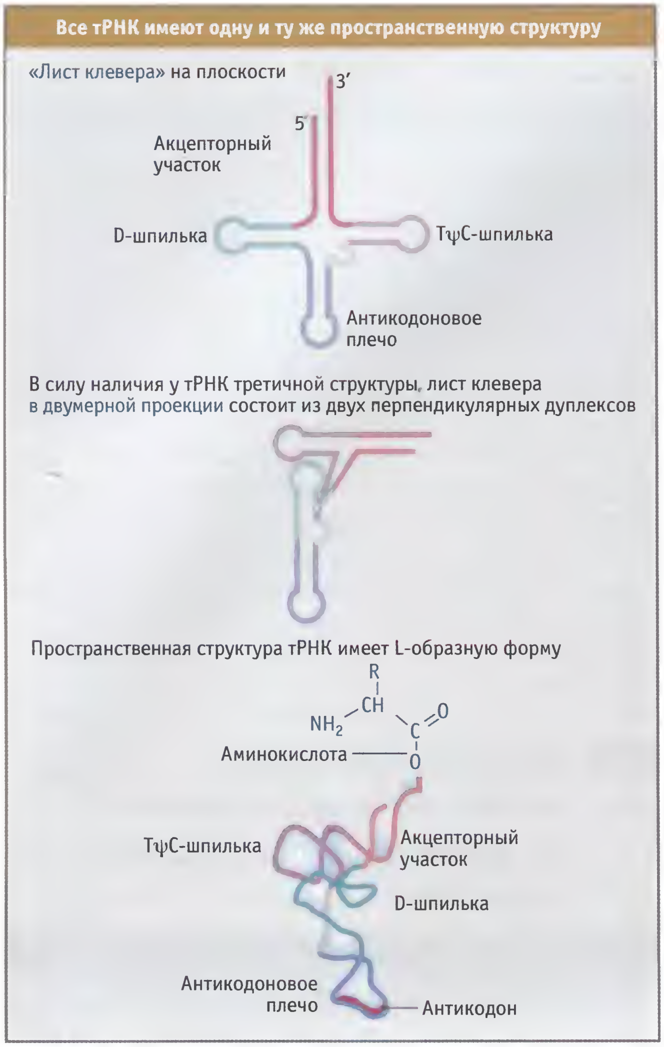


РИС. 7.6 Транспортная РНК уложена в компактную L-образную третичную структуру с аминокислотой и антикодоном на разных концах

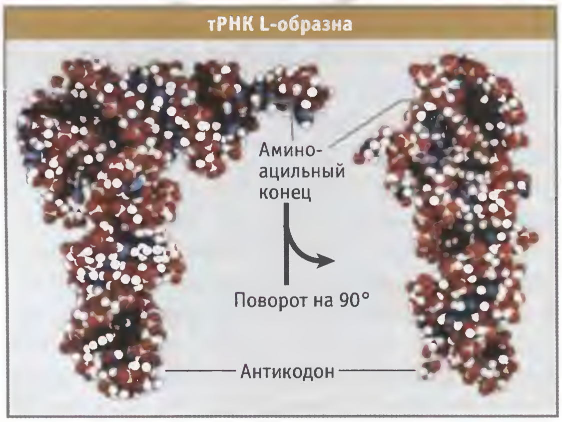


РИС. 7.7 Модель, созданная по методу заполнения пространства атомами, демонстрирует компактность третичной структуры тРНК фенилаланина (тРНК^{Phe}). Модель показана в двух ракурсах, повернутых друг относительно друга на 90°. Любезно предоставлено Sung-Hou Kim, University of California, Berkeley

водородные связи формируют, в частности, многие инвариантные или полуинвариантные основания, что объясняет их консервативность. Ни одна из этих связей не является универсальной, однако, вероятно, они определяют *генеральную* модель структуры тРНК.

Молекулярная модель структуры тРНК^{Phe} дрожжей приведена на РИС. 7.7. Изображение в левой части рисунка соответствует структуре, изображенной на нижней части рис. 7.6. В структуре других тРНК обнаруживаются отличия. Таким образом, несмотря на то, что все тРНК имеют сходную структуру, должен существовать способ распознавания отличий между ними. Так, например, угол между двумя осями тРНК^{Asp} чуть больше, и из-за этого молекула чуть более «открыта».

Рассматривая структуру и функции тРНК, можно прийти к следующему выводу: *участки молекулы, ответственные за выполнение конкретных функций, максимально удалены друг от друга.* Аминокислота расположена настолько далеко от антикодона, насколько это возможно, что согласуется с функцией, выполняемой тРНК при синтезе белка.

7.5 мРНК транслируется на рибосомах

Ключевые понятия:

- Рибосомы характеризуются константой седиментации (70S для бактериальных рибосом и 80S для рибосом эукариот).
- Рибосома состоит из большой субъединицы (50S и 60S для бактерий и эукариот соответственно) и малой субъединицы (30S или 40S).
- Рибосома создает условия, в которых аминоацил-тРНК могут добавлять аминокислоты к растущей полипептидной цепи в соответствии с последовательностью триплетов кодонов.
- Рибосома движется по мРНК от 5′- к 3′-концу.

Трансляция мРНК в полипептидную цепь катализируется **рибосомой**. Обычно рибосомы характеризуют (приблизительной) скоростью седиментации (или осаждения, измеряется в сведбергах; более высокое значение S означает большую скорость осаждения и большую массу). У бактерий константа седиментации рибосом около 70S, а у высших эукариот обычно около 80S.

Рибосома представляет собой компактную **рибонуклеопротеиновую** частицу, состоящую из двух субъединиц. Каждая субъединица состоит из РНК (включая одну очень длинную молекулу) и множества белков. Формирование рибосомы из двух субъединиц изображено на РИС. 7.8. *In vitro* две субъединицы диссоциируют с понижением концентрации ионов Mg^{2+}. В каждом случае **большая субъединица** примерно в два раза превышает по массе **малую субъединицу**. 70S рибосомы бактерий состоят из субъединиц с константами седиментации 50S и 30S, а цитоплазматические рибосомы эукариот (80S) имеют константы седи-



РИС. 7.8 Рибосома состоит из двух субъединиц

ментации 60S и 40S. Обе субъединицы работают вместе в составе рибосомы, однако каждая из них контролирует различные реакции в процессе синтеза белка.

Все рибосомы, расположенные в данном компартменте клетки, идентичны. *Взаимодействуя с различными мРНК, состоящими из кодирующих последовательностей, рибосомы осуществляют синтез различных белков.*

Рибосома создает условия, в которых осуществляется регуляция взаимодействия кодона мРНК и антикодона тРНК. При считывании генетического кода, записанного последовательностью триплетов, осуществляется синтез белка. Начало и конец синтезируемого полипептида соответствуют началу и концу кодирующего участка мРНК. *При продвижении рибосомы вдоль мРНК происходит сборка белка путем последовательного добавления аминокислот в направлении от N-конца к С-концу.*

Рибосома начинает трансляцию на 5′-конце кодирующего региона; продвигаясь к 3′-концу, она транслирует каждый триплет. Аминоацил-тРНК, соответствующая данному кодону, присоединяется к рибосоме и предоставляет свою аминокислоту для встраивания в полипептидную цепь. В каждый отдельно взятый момент рибосома может удерживать две аминоацил-тРНК, соответствующие следующим друг за другом кодонам. Это делает возможным формирование пептидной связи между аминокислотами. На каждом шаге растущая полипептидная цепь становится длиннее на одну аминокислоту.

7.6 К одной мРНК присоединяется множество рибосом

Ключевые понятия:

- мРНК может транслироваться несколькими рибосомами одновременно. Рибосомы, закрепленные на одной мРНК, находятся на разных стадиях синтеза.

Активные рибосомы с вновь синтезированными белками, выделенные в виде фракции, представляют собой комплексы, состоящие из мРНК, связанной с несколькими рибосомами. Это так называемая **полирибосома**, или **полисома**. 30S субъединица каждой рибосомы ассоциирована с мРНК, а 50S субъединица несет синтезированный белок. При этом тРНК расположена так, что она затрагивает обе субъединицы.

Продвигаясь по мРНК, каждая рибосома, входящая в состав полисомы, независимо синтезирует один полипептд. По сути, мРНК протянута через рибосому и каждый триплет нуклеотидов транслируется в аминокислоту. Таким образом, мРНК несет серии рибосом, формирующих белковый продукт, продвигаясь от 5′- к 3′-концу (РИС. 7.9). Полипептидную цепь, находящуюся в процессе синтеза, иногда называют насцентной (незрелой).



РИС. 7.9 Полирибосома содержит мРНК, которая транслируется сразу несколькими рибосомами, движущимися в направлении 5′–3′. Каждая рибосома несет две молекулы тРНК, одна из которых несет на себе синтезируемый полипептид. Вторая молекула тРНК несет следующую аминокислоту, которая должна быть добавлена к этому полипептиду



РИС. 7.10 На полисомах идет синтез белка. Фотография любезно предоставлена Alexander Rich, Massachusetts Institute of Technology



РИС. 7.11 Матричная РНК транслируется рибосомами. После завершения трансляции рибосомы подлежат повторному использованию

30% сухой массы бактерии имеет прямое отношение к экспрессии генов

Компонент	Составляемый им процент сухой клеточной массы	Число молекул на клетку	Число разновидностей этого компонента	Число копий каждой разновидности
Клеточная стенка	10	1	1	1
Мембрана	10	2	2	1
ДНК	1,5	1	1	1
мРНК	1	1 500	600	2–3
тРНК	3	200 000	60	>3 000
рРНК	16	38 000	2	19 000
Рибосомальные белки	9	10^6	52	19 000
Растворимые белки	46	$2{,}0 \times 10^6$	1 850	>1 000
Низкомолекулярные соединения	3	$7{,}5 \times 10^6$	800	

РИС. 7.12 Макромолекулярный состав *E. coli*

Примерно 30–35 последних добавленных в полипептидную цепь аминокислот закрыты от окружающей среды рибосомой. Вероятно, вся предшествующая часть полипептида выходит за пределы рибосомы и приобретает специфическую конформацию. Таким образом, белки отчасти могут принимать конформацию зрелой молекулы еще до того, как синтез завершен.

Классическая полисома показана на микрофотографии на РИС. 7.10. Глобин синтезируется набором из пяти рибосом, прикрепленных к мРНК (пентасомы). Рибосомы выявляются в виде сферических структур диаметром около 7 нм (70 Å), связанных нитью мРНК. Рибосомы расположены в различных местах мРНК. Те из них, что расположены на одном конце, только начали синтез, те же из них, что расположены на другом конце, находятся в конечной стадии синтеза полипептидной цепи.

Размер полисомы зависит от нескольких параметров. У бактерий полисомы очень большие. Они включают десятки рибосом, одновременно осуществляющих трансляцию. Отчасти размер полисом зависит от длины мРНК (которая у бактерий обычно кодирует несколько белков); кроме того, он зависит и от эффективности, с которой рибосомы прикрепляются к мРНК.

Полисомы, расположенные в цитоплазме эукариотических клеток, обычно меньше, чем у бактерий, однако здесь, как и в предыдущем случае, их размер зависит от длины мРНК (у эукариот обычно представляющей один белок) и от эффективности прикрепления рибосом. В среднем, мРНК эукариот в каждый момент несет примерно 8 рибосом.

На РИС. 7.11 изображен жизненный цикл рибосомы. Рибосомы временно изымаются из совокупности рибосомальных субъединиц, для того чтобы транслировать мРНК, а затем возвращаются назад. Число рибосом, расположенных на молекуле мРНК, варьирует как у бактерий, так и у эукариот. Оно зависит от статистических флуктуаций, определяемых длиной и эффективностью мРНК.

Доля компонентов, вовлеченных в синтез белка у интактной бактерии, приведена на РИС. 7.12. На долю ~20 000 рибосом приходится четверть клеточной массы. Каждая молекула тРНК представлена более чем 3000 копий, и все вместе они превосходят число рибосом почти в 10 раз. Большая часть из них представлена аминоацил-тРНК, т. е. они в любой момент готовы к использованию в синтезе белка. Число молекул мРНК довольно сложно оценить из-за их относительной нестабильности, однако можно предположить, что молекул, находящихся на разных стадиях синтеза, должно быть около 1500.

У бактерий около 600 различных типов мРНК. Причем на одну бактерию приходится две-три копии мРНК каждого типа. В среднем, каждая из них кодирует около 3 белков. Всего в клетках *E. coli* насчитывается 1850 различных глобулярных белков, и на одну бактерию приходится в среднем более 1000 копий каждого белка.

7.7 Жизненный цикл бактериальной мРНК

Ключевые понятия:

- **У бактерий транскрипция и трансляция происходят одновременно. Трансляция мРНК начинается еще до окончания ее синтеза.**
- **мРНК бактерий нестабильна. Период ее полураспада составляет всего несколько минут.**
- **Бактериальная мРНК может быть полицистронной с несколькими кодирующими регионами, представляющими разные гены.**

мРНК выполняет одинаковую функцию во всех клетках, однако в структуре и механизме синтеза прокариотической и эукариотической мРНК есть важные отличия.

В процессе продукции мРНК основное отличие касается мест локализации транскрипции и трансляции:

- У бактерий мРНК транскрибируется и транслируется в пределах одного клеточного компартмента; два этих процесса так тесно связаны, что происходят одновременно. Рибосомы прикрепляются к бактериальной мРНК еще до того, как заканчивается ее транскрипция. Таким образом, полисома оказывается прикрепленной к ДНК. Нестабильность бактериальной мРНК приводит к тому, что ее трансляция длится всего несколько минут.
- В эукариотической клетке синтез и созревание мРНК происходят исключительно в ядре. Только после того, как эти процессы завершены, мРНК экспортируется в цитоплазму, где и транслируется рибосомами. МРНК эукариот относительно стабильна, период ее полураспада в среднем составляет несколько часов.

На РИС. 7.13 показано, что у бактерий транскрипция и трансляция тесно взаимосвязаны. Транскрипция начинается, когда РНК-полимераза начинает продвигаться вдоль молекулы ДНК, синтезируя копию одной из цепей. Как только появляется мРНК, рибосома прикрепляется к 5′-концу и начинает трансляцию еще до того, как закончится синтез оставшейся части мРНК. Группа рибосом продвигается по мРНК в процессе ее синтеза. Транскрипция заканчивается с формированием 3′-конца мРНК. Рибосомы продолжают транслировать мРНК до тех пор, пока она сохраняет свою целостность, однако она довольно быстро деградирует в направлении от 5′- к 3′-концу. Жизнь одной молекулы бактериальной мРНК исчисляется единицами минут.

Скорость транскрипции и трансляции у бактерий примерно одинакова. При 37 °С скорость транскрипции составляет около 40 нуклеотидов в секунду, а скорость синтеза белка примерно 15 аминокислот в секунду. Таким образом, транскрипция и трансляция мРНК длиной 5000 нуклеотидов (что соответствует белку размером 180 кДа) займут примерно 2 минуты. При запуске экспрессии нового гена его мРНК обычно появляется в цитоплазме через 2,5 минуты. Соответствующий белок, вероятно, появится в течение следующих 30 секунд.

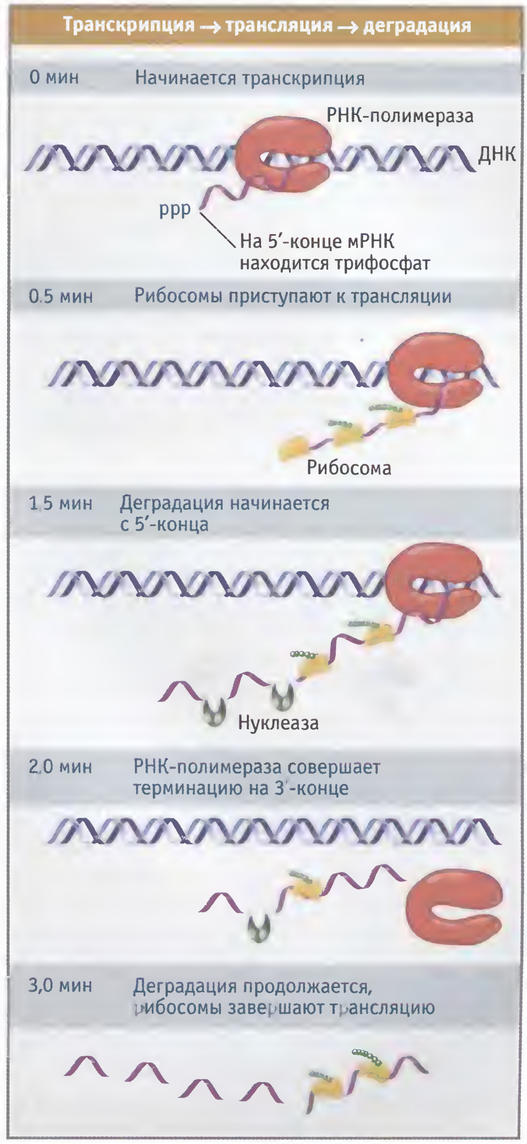


РИС. 7.13 У бактерий мРНК одновременно синтезируется, транслируется и деградирует

Трансляция у бактерий весьма эффективна, и большая часть мРНК транслируется значительным числом рибосом. Например, для *trp* мРНК каждую минуту транскрипция инициируется 15 раз, и каждая из 15 мРНК, вероятно, транслируется 30 рибосомами в промежутке между транскрипцией и деградацией.

Нестабильность большей части бактериальных мРНК поражает. Деградация мРНК начинается почти сразу же после трансляции и, скорее всего, начинается в течение минуты после начала транскрипции. 5′-конец мРНК начинает деградировать еще до того, как 3′-конец был синтезирован или транслирован. Деградация следует за последней в ряду рибосомой. Скорость деградации примерно в два раза ниже скорости транскрипции или трансляции.

Стабильность мРНК оказывает основное влияние на количество продуцируемого белка. Стабильность

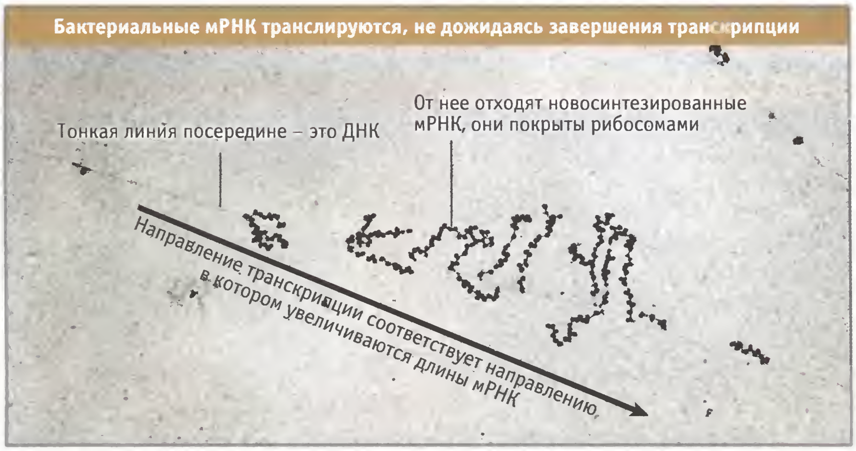


РИС. 7.14 Единицы транскрипции у бактерий можно увидеть. Фотография любезно предоставлена Oscar Miller

обычно выражается скоростью полураспада. мРНК, представляющая любой отдельно взятый ген, характеризуется определенной скоростью полураспада, однако в среднем у бактерий она составляет около 2 минут.

Вся эта череда событий, конечно, возможна только лишь потому, что транскрипция, трансляция, и деградация идут в одном направлении. Экспрессия гена была «поймана с поличным» на микрофотографии, представленной на РИС. 7.14. На этой (неизвестной) транскрипционной единице происходит синтез нескольких мРНК, и каждая из них несет несколько рибосом, осуществляющих трансляцию. (Момент, пойманный на фотографии, соответствует схеме, изображенной на второй сверху иллюстрации рис. 7.13.) РНК, синтез которой еще не завершен, часто называют **незрелой РНК**.

мРНК бактерий сильно отличаются по числу кодируемых ими белков. Некоторые мРНК представляют только один ген и называются **моноцистронными**. Последовательности других (подавляющего большинства) кодируют несколько белков и называются **полицистронными**. В этом случае одна мРНК транскрибируется с группы прилегающих друг к другу генов. (Подобный кластер генов формирует оперон, представляющий собой одну генетическую единицу; см. гл. 12 «Оперон».)

мРНК состоит из участков двух типов. Кодирующий регион состоит из серии кодонов, представляющих аминокислотную последовательность белка и чаще всего начинается с кодона AUG и заканчивается стоп-кодоном. МРНК всегда длиннее кодирующего региона, поскольку по обоим концам она несет дополнительные участки. Последовательность на 5′-конце, предшествующую началу кодирующего региона, называют лидирующей или **5′-UTR** (5′-нетранслируемым регионом). Дополнительный участок, следующий за сигналом терминации, формирует 3′-конец, и его называют трейлером или **3′-UTR** (3′-нетранслируемым регионом). Несмотря на то что эти последовательности являются частью транскрипционной единицы, они не кодируют белок.

Кроме того, полицистронная мРНК содержит **внутрицистронные регионы** (РИС. 7.15). Они сильно отличаются по размеру: в бактериальных мРНК они иногда достигают 30 нуклеотидов в длину (а у фагов встречаются еще более длинные), а иногда они совсем невелики и состоят из одного-двух нуклеотидов, отделяя стоп-кодон одного белка от старт-кодона другого. В редких случаях два гена даже перекрываются таким образом, что последний нуклеотид одного кодирующего региона является первым нуклеотидом следующего.



РИС. 7.15 В бактериальной мРНК есть транслируемые и нетранслируемые области. Каждая кодирующая область имеет свои собственные сигналы инициации и терминации. Характерным является наличие нескольких кодирующих областей в одной мРНК

Число рибосом, вовлеченных в трансляцию данного цистрона, зависит от эффективности его сайта инициации. Сайт инициации первого цистрона становится доступным сразу после того, как синтезируется 5′-конец мРНК. Как же транслируются последующие цистроны? Транслируются кодирующие регионы полицистронной мРНК независимо или же эти участки экспрессируются взаимосвязанно? Одинаков механизм инициации для всех цистронов или же он отличается у первого и всех последующих цистронов?

Трансляция мРНК бактерий идет последовательно цистрон за цистроном. В тот момент, когда рибосома присоединяется к первому кодирующему региону, следующий за ним регион еще даже не транскрибирован. К тому времени как открывается сайт связывания для второй рибосомы, трансляция первого цистрона только заканчивается. Обычно рибосомы заканчивают трансляцию в конце первого цистрона (и диссоциируют на субъединицы), а новая рибосома независимо собирается в начале следующего кодирующего региона. (Процессы инициации и терминации более подробно рассмотрены в гл. 8 «Синтез белка».)

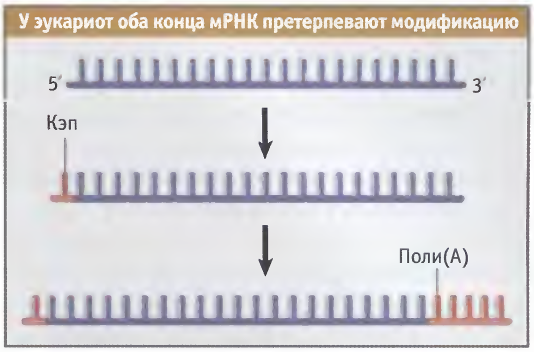


РИС. 7.16 Эукариотическая мРНК подлежит модификации: к 5′-концу добавляется кэп, к 3′-концу добавляется «хвост» из поли(А)

7.8 мРНК эукариот претерпевает изменения в процессе или после транскрипции

Ключевые понятия:

- У эукариот мРНК-транскрипт претерпевает изменения в ядре (во время или вскоре после транскрипции).
- Изменения включают добавление метилированного кэпа к 5′-концу и последовательности поли(А) к 3′-концу.
- мРНК переносится из ядра в цитоплазму только после завершения всех модификаций.

У эукариот синтез мРНК включает дополнительные стадии. Транскрипция начинается как обычно: с формирования 5′-трифосфатного конца. Однако 3′-конец формируется не за счет остановки транскрипции в определенном месте, а за счет разрезания транскрипта. Для удаления из РНК интронов, синтезированных с прерывистых генов, необходим сплайсинг, в результате которого формируются маленькие мРНК, содержащие непрерывные кодирующие последовательности.

На **РИС. 7.16** видно, что транскрипт модифицирован с обеих сторон путем добавления дополнительных нуклеотидов (для этого требуются дополнительные ферменты). 5′-конец мРНК практически сразу после формирования модифицируется добавлением «кэпа». Это приводит к развороту трифосфата исходного транскрипта в обратном направлении (3′–5′), и конец, таким образом, запечатывается. 3′-конец модифицируется добавлением серии аденинов [полиадениловая кислота или **поли(А)**] сразу после разрезания. Только после завершения модификации и процессинга мРНК может быть экспортирована из ядра в цитоплазму. В среднем, до выхода в цитоплазму проходит

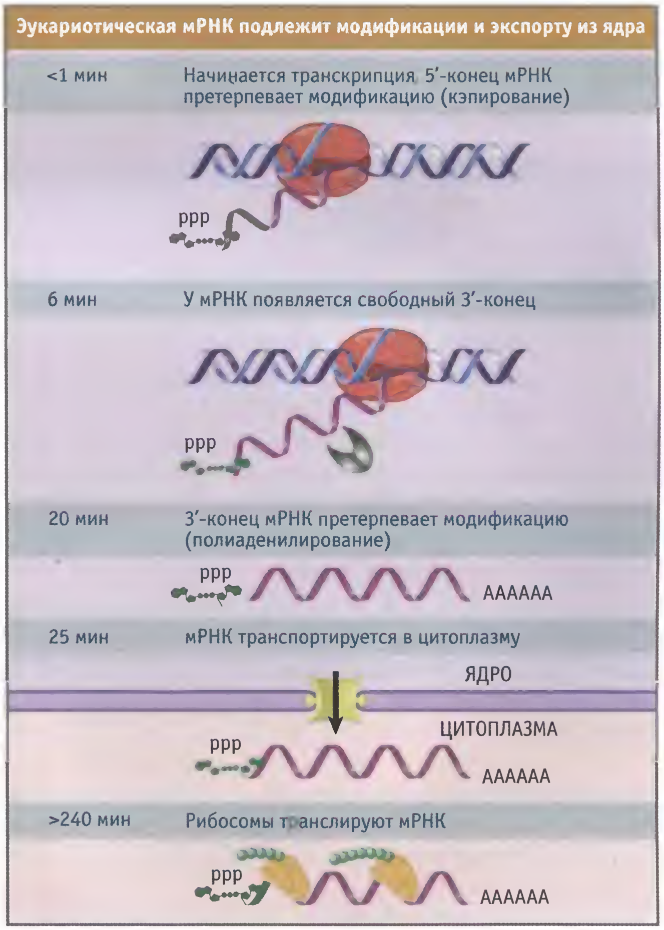


РИС. 7.17 Путь мРНК в клетках животных очень сложен и включает в себя транскрипцию, модификацию, процессинг, транспорт из ядра в цитоплазму и, наконец, трансляцию

примерно 20 минут. Как только мРНК оказалась в цитоплазме, рибосомы распознают и транслируют ее.

На **РИС. 7.17** показано, что жизненный цикл мРНК эукариот более продолжителен, чем у бактерий. Скорость транскрипции в клетках животных примерно такая же, как

и у бактерий, — около 40 нуклеотидов в секунду. Многие эукариотические гены очень велики; транскрипция гена длиной 10 000 пн занимает около 5 минут. Транскрипция мРНК не завершается после отделения ферментов от ДНК и может продолжаться за пределами гена. Последовательная серия событий формирует 3′-конец мРНК путем разрезания и добавления к нему поли(А).

мРНК эукариот составляет лишь малую долю от общей массы клеточной РНК (около 3%). У дрожжей период полураспада относительно невелик и занимает 1–60 мин. У высших эукариот стабильность мРНК значительно увеличивается; период полураспада мРНК в клетках животных варьирует от 4 до 24 ч. Полисомы эукариот значительно более стабильны. Важную роль в увеличении стабильности мРНК играет модификация ее концов.

7.9 У эукариот 5′-конец мРНК кэпирован

Ключевые понятия:

- 5′-кэп формируется путем добавления G к крайнему нуклеотиду транскрипта посредством 5′-5′ связи. К основанию или рибозе нового концевого гуанидина добавляется от одной до трех метильных групп.

Транскрипция начинается с нуклеозидтрифосфата (обычно пурина: A или G). Первый нуклеотид сохраняет свою 5′-трифосфатную группу и формирует обычную фосфодиэфирную связь от своей 3′-позиции к 5′-позиции следующего нуклеотида. Исходная последовательность транскрипта может быть изображена следующим образом:

$$5'ppp^{A}/_{G}pNpNpNp...$$

Если обработать зрелую мРНК *in vitro* ферментами, разрушающими ее до отдельных нуклеотидов, на 5′-конце не окажется ожидаемого трифосфата. Вместо этого на 5′-конце расположены два нуклеотида, соединенные 5′–′5-трифосфатной связью и несущие метильные группы. Концевой нуклеотид — это всегда гуанин, который добавляется к исходной молекуле РНК после начала транскрипции.

Добавление G к 5′-концу катализирует фермент гуанилилтрансфераза. Присоединение происходит практически сразу после начала транскрипции, настолько быстро, что исходный 5′-трифосфат ядерной РНК можно обнаружить лишь в следовых количествах. В целом реакцию можно представить в виде конденсации между ГТФ и исходным трифосфатом, расположенным на 5′-конце РНК:

$$\overset{5'}{Gppp} + \overset{5'}{pppApNpNp...}$$
$$\downarrow$$
$$\overset{5'-5'}{GpppApNpNp...} + pp + p$$

Новый остаток G добавляется к концу РНК в противоположной всем остальным нуклеотидам ориентации.

Эта структура называется **кэп**. Она служит субстратом для нескольких реакций метилирования. На РИС. 7.18 показана полная структура кэпа после добав-

5′-конец эукариотической мРНК несет метилированный кэп

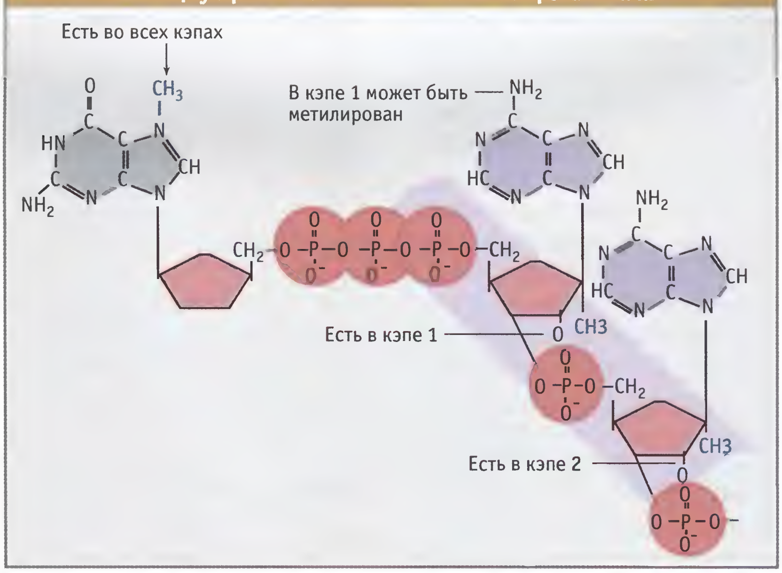


РИС. 7.18 Кэп блокирует 5′-конец мРНК; он может быть метилирован по нескольким положениям

ления всех возможных метильных групп. По количеству метилирований кэпы делят на несколько типов:

- У эукариот в результате первого метилирования метильная группа добавляется в 7-ю позицию концевого гуанина. Этот тип метилирования встречается у всех эукариот. Кэп, несущий только эту группу, известен как **кэп 0**. Данная реакция берет начало от одноклеточных эукариот. Ферментом, ответственным за одноклеточную модификацию, является гуанин-7-метилтрансфераза.
- На следующем этапе еще одна метильная группа добавляется в положение 2′-O предпоследнего нуклеотида (который на самом деле является первым нуктеотидом в немодифицированном транскрипте). Эту реакцию катализирует другой фермент (2′-O-метил-трансфераза). Кэп с двумя метильными группами называют **кэп 1**. Это преобладающий тип кэпов у всех эукариот (за исключением одноклеточных).
- В небольшом проценте случаев ко второму нуклеотиду добавляется еще одна метильная группа. Это происходит только в том случае, если в данной позиции стоит аденин. Метильная группа в этом случае добавляется в положение N^6. Ответственный за это фермент действует только на аденозиновый субстрат, который уже несет метильную группу в положении 2′-O.
- У некоторых видов метильная группа добавляется к третьему нуклеотиду кэпированной РНК. Субстратом для этой реакции является кэп1 мРНК, которая уже несет две метильные группы. Модификация третьего нуклеотида — это всегда 2′-O-метилирование рибозы. Этот тип кэпов называют **кэп 2**. Обычно этот тип составляет менее 10–15% от общей кэпированной популяции транскриптов.

В пуле эукариотической мРНК все молекулы кэпированы. Относительная представленность кэпов различных типов является характерной чертой данного организма. Нам неизвестно, является структура определенной мРНК постоянной или же она может иметь разные типы кэпов.

В мРНК высших эукариот, помимо метилирования кэпов, с незначительной частотой происходят внутренние метилирования. В результате таких реакций в мРНК с частотой одна модификация на 1000 нуклеотидов формируются N^6 метиладениновые остатки. В типичной для высших эукариот мРНК встречается один или два метиладенина, однако их наличие не является обязательным; некоторые молекулы не имеют их вовсе.

7.10 3′-конец полиаденилирован

Ключевые понятия:

- К ядерному транскрипту после транскрипции добавляется участок поли(A) длиной около 200 нуклеотидов.
- Поли(A) связывается специфическим белком (PABP).
- Поли(A) предохраняет мРНК от деградации.

Последовательность остатков A на 3′-конце мРНК называют поли(A) хвостом; мРНК, несущую такой хвост, обозначают как **поли(A)$^+$**.

Последовательность поли(A) не закодирована в ДНК, а добавляется к РНК в ядре после транскрипции. Добавление поли(A) катализирует фермент **поли(A)-полимераза**, который добавляет около 200 остатков A к свободному 3′-OH концу мРНК. Участок поли(A) как ядерной, так и мРНК связан с **поли(A)-связывающим белком** (PABP, poly(A)-binding protein). Похожие формы этого белка встречаются у многих эукариот. Один мономер PABP размером около 70 кДа связывает каждые 10–20 нуклеотидов поли(A) хвоста. Общей особенностью большинства или всех эукариот является то, что 3′-конец их мРНК состоит из последовательности поли(A), связанной с большим количеством белка. Добавление поли(A) является частью сложной реакции, в которой формируется и модифицируется многими ферментами 3′-конец мРНК (см. разд. 26.19 «3′-концы мРНК формируются в результате разрезания и полиаденилирования»).

Связывание PABP и фактора инициации eIF4G формирует закрытую петлю, в которой 5′- и 3′-концы мРНК находятся в общем белковом комплексе (см. рис. 8.20 в разд. 8.9 «Эукариоты используют комплекс, состоящий из множества факторов ассоциации»). Формирование этого комплекса может отвечать за влияние поли(A) на свойства мРНК. Способность поли(A) защищать мРНК от деградации зависит от связывания с PABP.

Удаление поли(A) ингибирует инициацию трансляции *in vitro*. У дрожжей тот же эффект наблюдается при уменьшении концентрации PABP. Эти эффекты могут зависеть от связывания PABP и комплекса инициации на 5′-конце мРНК. В раннем эмбриогенезе существует множество примеров, когда полиаденилирование отдельной мРНК связано с ее трансляцией. В одних случаях мРНК хранятся в неполиаденилированной форме, а поли(A) добавляется, когда возникает необходимость их трансляции; в других случаях поли(A)$^+$ мРНК деаденилируются, и их трансляция прекращается.

Существование поли(A) имеет важное практическое значение для исследователей. Участок поли(A) мРНК может связываться с олиго(U) или олиго(dT), что может быть использовано при выделении поли(A)$^+$ мРНК. Самая удобная методика состоит в закреплении олигонуклеотидов (U или dT) на твердой подложке. Тогда при пропускании РНК через колонку на сорбент осядет только поли(A)$^+$ мРНК (рис. 7.19). Выделить ее можно, обработав колонку раствором, разрушающим сформировавшиеся водородные связи.

Единственное неудобство этого метода состоит в том, что выделяется вся РНК, содержащая поли(A). Если мы берем из клетки всю РНК, то при выделении получим как ядерную, так и цитоплазматическую поли(A)$^+$ РНК. При использовании препаратов полисом (стандартная процедура) большая часть изолированной поли(A)$^+$ РНК будет активной мРНК.

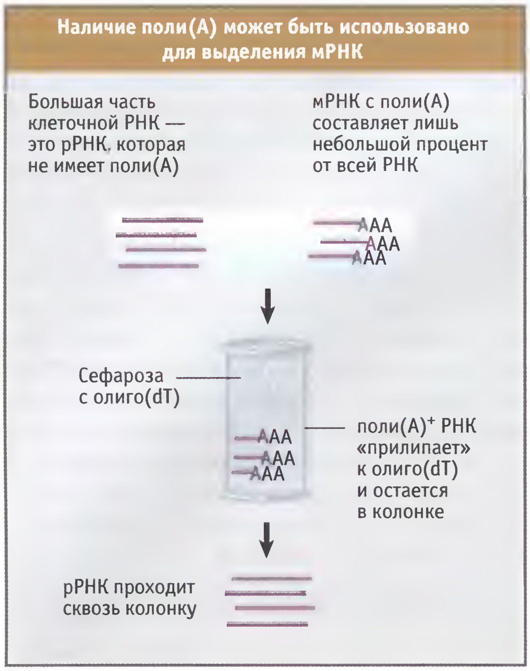


РИС. 7.19 Поли(А)$^+$ РНК можно отделить от остальной РНК посредством сорбции на сефарозе с олиго(dT)

Однако помимо мРНК полисом мы также выделим частицы рибонуклеопротеина из цитозоля, которые содержат нетранслируемую поли(А)$^+$ мРНК. Эта РНК, вероятно, хранится для использования в другое время. Таким образом, выделенная тотальная поли(А)$^+$ мРНК не соответствует популяции активной мРНК.

Для исследования различных мРНК используют «клонирование». Это процедура, в ходе которой мРНК копируется для производства комплементарной цепи ДНК (**кДНК**). Затем кДНК используют в качестве матрицы для синтеза второй цепи ДНК (идентичной исходной последовательности мРНК). Продуктом этих реакций является двухцепочечная ДНК, соответствующая последовательности мРНК. Такую ДНК можно получить в больших количествах.

Доступность клонированых ДНК упрощает процесс выделения соответствующей мРНК гибридизационными методами. Даже мРНК, представленная в клетке несколькими копиями, может быть выделена подобным образом. На самом деле напрямую, без клонирования, можно выделить только мРНК, представленную в достаточно большом количестве.

Почти все мРНК в клетке имеют поли(А). Важным исключением являются мРНК, кодирующие гистоны (основной структурный компонент хромосом). Практически только мРНК гистонов составляют фракцию **поли(А)**$^-$. Не совсем понятно, почему у них отсутствует поли(А)-хвост? Кроме того, трудно найти функцию, в связи с выполнением которой это было бы необходимо.

7.11 В деградацию мРНК бактерий вовлечено множество ферментов

Ключевые понятия:

- мРНК бактерий деградирует в направлении от 5′- к 3′-концу.
- Деградация — это результат серии нескольких расщеплений молекулы эндонуклеазами и последующего разрушения образовавшихся фрагментов экзонуклеазами в направлении от 3′- к 5′-концу

мРНК бактерий подвергается постоянной деградации, осуществляемой набором эндо- и экзонуклеаз. Эндонуклеазы разрезают молекулу РНК по внутренним сайтам. Экзонуклеазы вовлечены в реакции «тримминга», в ходе которых нуклеотиды удаляются с конца молекулы. У бактерий экзонуклеазы, «работающие» с одноцепочечной ДНК, продвигаются вдоль молекулы в направлении от 3′-конца.

Механизм согласованного действия эндо- и экзонуклеаз, участвующих в деградации мРНК, изображен на **РИС. 7.20**. Запуск деградации бактериальной мРНК осуществляют эндонуклеазы. В результате разрезания молекулы эндонуклеазами может сформироваться несколько 3′-концов. Молекула разрушается от 5′- к 3′-концу

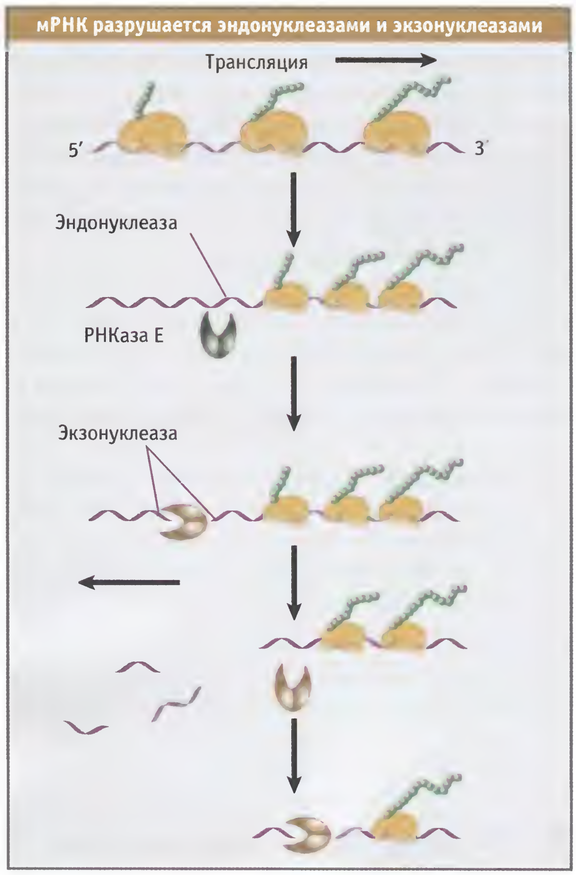


РИС. 7.20 Деградация бактериальной мРНК — это двуступенчатый процесс. Позади рибосом, транслирующих мРНК, проводится серия эндонуклеолитических разрезов (серия идет от 5′-конца к 3′-концу мРНК). Получающиеся при этом небольшие фрагменты РНК подлежат деградации экзонуклеазами в направлении 3′–5′

(это определяют по потере способности синтезировать белки). Вероятно, это происходит в результате серии разрезаний участка молекулы, расположенного за последней рибосомой. Разрушение фрагментов мРНК, возникших в результате действия эндонуклеаз, продолжают экзонуклеазы, которые расщепляют молекулу до нуклеотидов, продвигаясь от 3′-конца к 5′-концу фрагмента (в направлении, противоположном транскрипции). Эндонуклеазы расщепляют молекулу на фрагменты, обладающие различной восприимчивостью к действию экзонуклеаз. Участок, формирующий вторичную структуру в пределах мРНК, может послужить препятствием для экзонуклеазы и «прикрыть» последовательность с 5′-конца. Таким образом, стабильность каждой мРНК определяется предрасположенностью ее последовательности к воздействию эндо- и экзонуклеаз.

У *E. coli* около 12 рибонуклеаз. Штаммы, мутантные по генам эндорибонуклеаз (за исключением рибонуклеазы I, которая не оказывает эффекта), жизнеспособны, однако накапливают непроцессированные предшественники рРНК и тРНК. Мутанты по экзонуклеазам часто фенотипически не отличаются от штаммов дикого типа, то есть в отсутствие одного из ферментов другой может выполнять его функцию. Мутанты по нескольким ферментам часто нежизнеспособны.

Ключевым ферментом в процессе запуска деградации мРНК является РНКаза Е. Вероятно, этот фермент в большинстве случаев осуществляет первое разрезание. Мутантные бактерии, несущие дефектную рибонуклеазу Е, отличаются повышенной (в 2 или 3 раза) стабильностью мРНК. Но это не единственная функция РНКазы. Изначально РНКаза Е была определена как фермент, ответственный за 5′-процессинг первичного транскрипта рРНК путем его специфического разрезания.

Процесс деградации может поддерживаться каталитической активностью полиферментного комплекса, иногда называемого **деградосомой**. В состав деградосомы входят: рибонуклеаза Е, полинуклеотидфосфатаза и хеликаза. РНКаза Е выполняет двойную функцию. Ее N-концевой домен обладает эндонуклеазной активностью, в то время как С-концевой домен выполняет функцию каркаса, удерживающего остальные компоненты деградосомы. Функция хеликазы заключается в раскручивании РНК-субстрата для его последующего представления полинуклеотидфосфатазе. В соответствии с этой моделью РНКаза Е совершает первый разрез и затем передает фрагменты остальным компонентам комплекса, осуществляющим дальнейший процессинг.

У бактерий полиаденилирование может играть определенную роль в деградации некоторых мРНК. У *E. coli* поли(А)-полимераза ассоциируется с рибосомами и добавляет короткие (10–40 нуклеотидов) фрагменты поли(А), по крайней мере к некоторым мРНК. Тройные мутации, приводящие к удалению поли(А)-полимеразы, рибонуклеазы Е и полинуклеотидфосфатазы (полинуклеотидфосфатаза является 3′–5′-экзонуклеазой), оказывают значительное влияние на стабильность мРНК. (По отдельности или в двух генах они оказывают слабый эффект.) Поли(А)-полимераза может формировать поли(А)-хвост, выполняющий роль сайта связывания для нуклеаз. Таким образом, функция поли(А) у бактерий может отличаться от таковой у эукариот.

7.12 Стабильность мРНК зависит от ее структуры и последовательности

Ключевые понятия:

- Модификации на концах мРНК защищают ее от разрушения экзонуклеазами.
- Некоторые последовательности мРНК могут стабилизировать или дестабилизировать молекулу.
- Дестабилизация может быть спровоцирована потерей поли(А).

Основные особенности мРНК, влияющие на ее стабильность, приведены на РИС. 7.21. Важную роль здесь играют и структура, и последовательность. 5′- и 3′-кон-

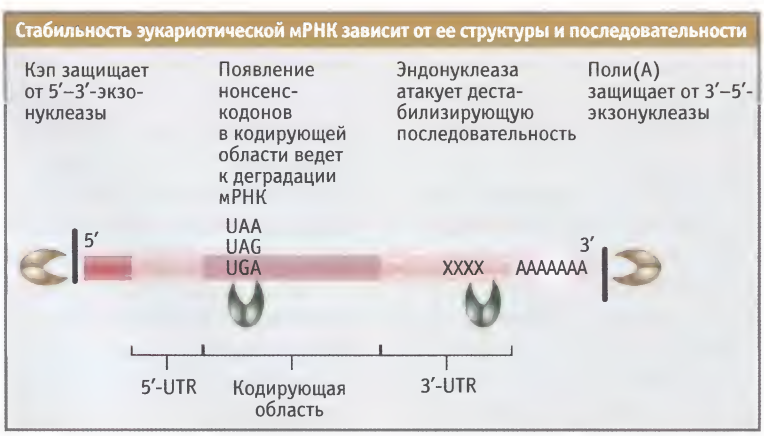


РИС. 7.21 Концевые модификации защищают мРНК от деградации. Определенные последовательности мРНК могут активировать системы деградации

цевые структуры защищают молекулу от деградации, а специфические последовательности в пределах мРНК могут как служить мишенью для запуска деградации, так и защищать от нее:

- Изменения на 5'- и 3'-концах мРНК играют важную роль в предотвращении экзонуклеазных атак. Кэп защищает 5'-конец от воздействия 5'→3'-экзонуклеаз, а поли(А) защищает 3'-конец от атак со стороны 3'→5'-экзонуклеаз.
- Определенные элементы последовательности в пределах мРНК могут принимать участие как в ее стабилизации, так и в дестабилизации. Дестабилизирующие элементы чаще всего расположены в пределах 3'-нетранслируемого региона. Наличие такого элемента укорачивает время существования мРНК.
- Мутации, формирующие стоп-кодоны в пределах кодирующего участка, запускают систему, разрушающую мРНК (см. разд 7.14 «Нонсенс-мутации запускают систему надзора»).

Дестабилизирующие элементы были обнаружены в нескольких мРНК дрожжей, однако на сегодняшний день мы не располагаем данными о их распространенности, а также о том, как они дестабилизируют мРНК. Их влияние может быть не прямым (как если бы они служили мишенью для эндонуклеаз), а опосредованным (например, через стимуляцию деаденилирования). Критерием, определяющим дестабилизирующий элемент, является деградация молекулы мРНК при встраивании в нее такой последовательности. Удаление дестабилизирующего элемента из мРНК может никак не повлиять на ее стабильность, что, вероятно, объясняется наличием других дестабилизирующих элементов в последовательности.

Общей особенностью некоторых нестабильных мРНК является наличие AU-богатых последовательностей протяженностью около 50 нуклеотидов (часто называемых ARE-элементами), расположенных в регионе 3'-конца. Консенсусной последовательностью ARE-элементов является пентануклеотид AUUUA, повторяющийся несколько раз. На РИС. 7.22 показано, что запуск процесса дестабилизации посредством ARE-элементов осуществляется в 2 этапа: сначала мРНК деаденилируется, а затем разрушается. Вероятнее всего, деаденилирование необходимо для сбрасывания поли(А)-связывающего белка, присутствие которого стабилизирует 3'-регион (см. разд. 7.13 «В деградацию мРНК вовлечено множество механизмов»).

В некоторых случаях мРНК может быть стабилизирована за счет специфического ингибирования функции дестабилизирующего элемента. Так, мРНК трансферрина содержит последовательность, называемую IRE (Iron-responsive element), которая контролирует ответ мРНК на изменение концентрации железа. IRE расположен в 3'-нетранслируемом регионе и несет структуры типа шпилька–петля, связывающиеся с белком, афинность которого к мРНК опосредована железом. На РИС. 7.23 показано, что связывание белка с IRE стабилизирует мРНК за счет ингибирования функции

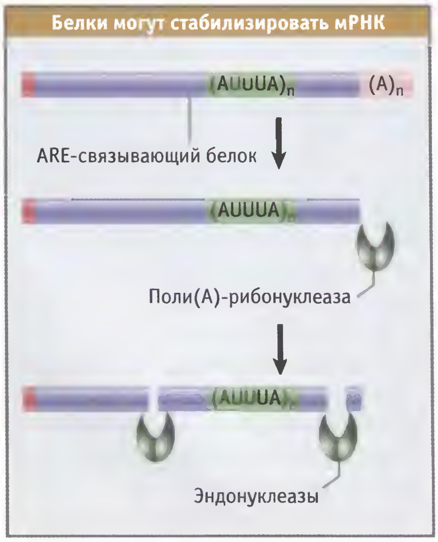


РИС. 7.22 Элемент ARE в 3'-некодирующей области инициирует деградацию мРНК



РИС. 7.23 Элемент IRE, находящийся в 3'-некодирующей области мРНК, участвует в контроле ее стабильности

расположенных поблизости (неизвестных) дестабилизирующих последовательностей. Видимо, повышение стабильности за счет ингибирования функции дестабилизирующих последовательностей является общей моделью стабилизации мРНК.

Декэпирование и деградация происходят с участием крупных белковых комплексов, которые могут быть локализованы в виде отдельных цитоплазматических единиц, назывемых Р-телами. В области Р-тел локализованы белки, вовлеченные в деградацию и метаболизм мРНК. По сути, Р-тела могут осуществлять уничтожение любой мРНК, не участвующей в процессе трансляции.

7.13 В деградацию мРНК вовлечены ферменты с различными активностями

Ключевые понятия:

- Деградация мРНК дрожжей требует удаления 5′-кэпа и 3′-поли(А).
- Один из путей деградации мРНК дрожжей основан на ее расщеплении экзонуклеазами рестрикции в направлении от 5′- к 3′-концу.
- Другой путь деградации основан на действии комплекса из нескольких экзонуклеаз, работающих в направлении от 3′- к 5′-концу.
- Деаденилаза клеток животных может напрямую связываться с 5′-кэпом.

Деградацию мРНК у дрожжей изучена достаточно хорошо. В ее основе лежат два пути. Оба начинаются с удаления поли(А)-хвоста. Эту реакцию катализирует специфическая деаденилаза, которая, вероятно, осуществляет свои функции, находясь в составе крупного белкового комплекса. (Каталитической субъединицей у дрожжей является экзонуклеаза Ccr4; у позвоночных ту же функцию выполняет экзонуклеаза PARN, родственная РНКазе D.) Фермент осуществляет деградацию мРНК. Реакция включает последовательное отщепление нуклеотидов и продолжается до тех пор, пока вся молекула не будет разрушена.

Основной путь деградации изображен на РИС. 7.24. Деаденилирование 3′-конца запускает декэпирование 5′-конца. Связь этих событий осуществляется посредством белка PABP. Находясь на поли(А), он препятствует связыванию декэпирующего фермента с 5′-концом молекулы. При уменьшении поли(А) хвоста до 10–15 нуклеотидов PABP отсоединяется от молекулы. Декэпирование происходит за счет отрезания метилированного нуклеотида от 5′-конца с высвобождением монофосфата: $m^7GpppX... \rightarrow m^7GDP + pX...$ Для фермента необходимо наличие 7-метильной группы.

Каждый из концов мРНК может оказывать влияние на события, происходящие на другом конце. Это объясняется тем, что оба они удерживаются вместе факторами, вовлеченными в синтез белка (см. разд. 8.9 «В эукариотических клетках используется комплекс из нескольких факторов инициации»). Влияние, оказываемое PABP на декэпирование, позволяет 3′-концу стабилизировать 5′-конец. Кроме того, процесс деградации 5′-конца связан с наличием специфической структуры на 5′-конце. Деаденилаза напрямую связывается с 5′-кэпом, и это взаимодействие запускает эндонуклеазную атаку на поли(А)-хвост.

Зачем же нужна связь между событиями, происходящими на разных концах мРНК? Вероятно, это позволяет исключить возможность повторного формирования активной мРНК.

Удаление кэпа запускает 5′–3′ путь деградации, характеризующийся быстрым разрушением мРНК с 5′-конца за счет действия 5′–3′ экзонуклеазы XRN1. Фермент,

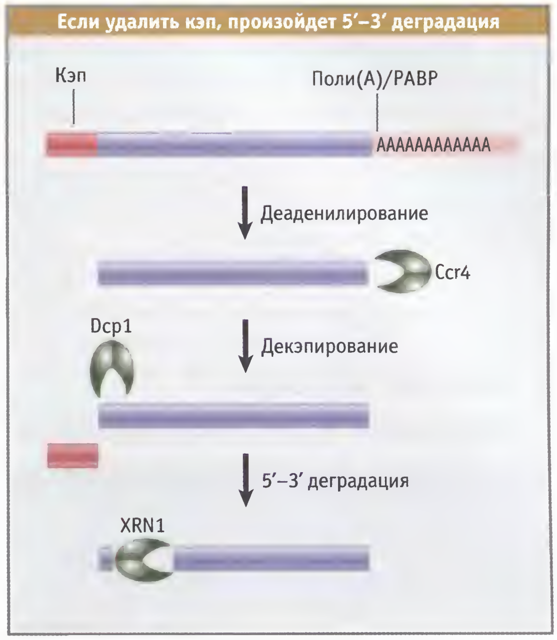


РИС. 7.24 Деаденилирование способствует декэпированию, которое, в свою очередь, приводит к экзонуклеолитическому расщеплению мРНК с 5′-конца

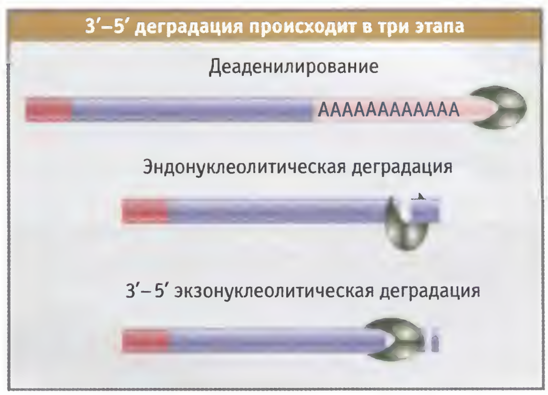


РИС. 7.25 Деаденилирование может напрямую вести к эндонуклеолитическому расщеплению и экзонуклеолитическому расщеплению с 3′-конца мРНК

осуществляющий декэпирование, сконцентрирован в отдельных цитоплазматических компартментах «процессирующими комплексами», осуществляющими деаденилирование, декэпирование и деградацию мРНК.

При деградации мРНК дрожжей по второму пути разрушение идет за счет 3′–5′ экзонуклеазной активности экзосомы — комплекса, включающего более 9 экзонуклеаз. Кроме того, экзосома вовлечена в процессинг предшественников рРНК. Сборка экзонуклеаз в экзосому позволяет контролировать 3′–5′ экзонуклеазную активность. Экзосома также может разрушать фрагменты мРНК, сформированные в результате расщепления молекулы эндонуклеазами. На РИС. 7.25 по-

казано, что 3′–5′ путь деградации может использовать как эндо-, так и экзонуклеазные активности ферментов. Помимо цитоплазмы, экзосома часто встречается в ядре, где она разрушает предшественников мРНК, не прошедших сплайсинг.

Дрожжевые клетки, в которых инактивирован экзонуклеазный путь, разрушают мРНК медленнее. Утрата обоих путей приводит к гибели клетки.

7.14 Нонсенс-мутации запускают «систему надзора»

Ключевые понятия:

- Нонсенс-мутации являются причиной деградации мРНК.
- Гены, кодирующие систему деградации, были найдены у дрожжей и червей.

Еще один путь деградации — **опосредованное нонсенс-мутациями разрушение мРНК** (РИС. 7.26). Вставка нонсенс-мутации часто ведет к ускорению деградации мРНК. Поскольку эти мутации формируют стоп-кодон, можно предположить, что деградация происходит в цитоплазме. Этот механизм представляет собой своеобразный контроль качества или **«систему надзора»**, ответственную за удаление нефункциональных мРНК.

«Система надзора» лучше всего была изучена у дрожжей и *C. elegans*, однако она может играть важную роль

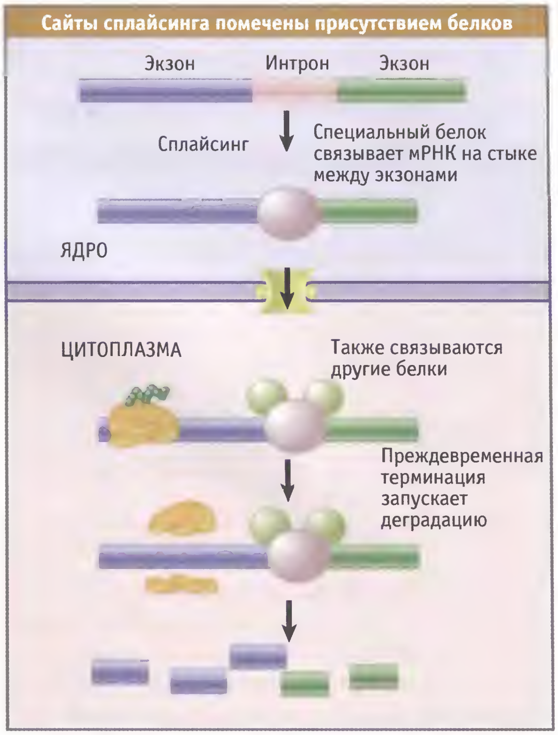


РИС. 7.27 В «системе надзора» должен быть хотя бы один ядерный белок, который связывал бы зрелую мРНК по «горячим точкам» прошедшего сплайсинга. Другие белки могут вступать с ним в комплекс в ядре или уже в цитоплазме. Согласно принципу работы «системы», эти белки побуждаются к деградации мРНК в том случае, если рибосомы терминируют трансляцию преждевременно

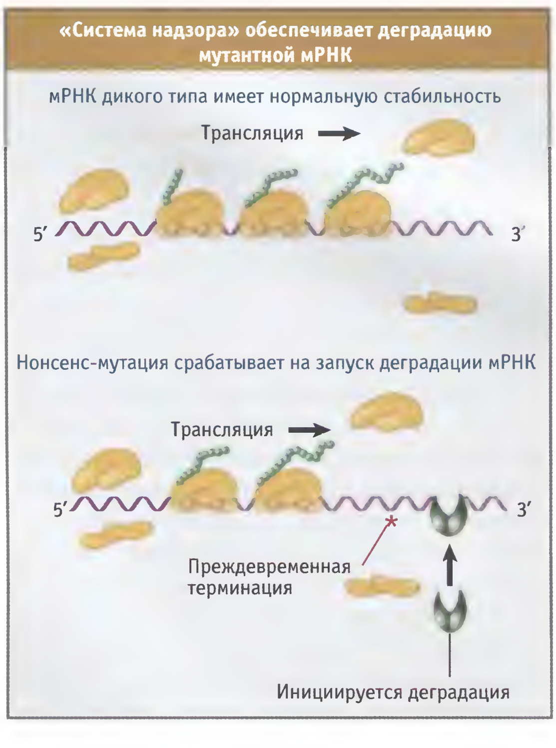


РИС. 7.26 Нонсенс-мутации могут приводить к деградации мРНК

и в клетках животных. Так, например, при формировании иммуноглобулинов и Т-клеточных рецепторов модификация генов происходит за счет соматической рекомбинации и мутации (см. гл. 23 «Многообразие генов иммунной системы»). Это приводит к формированию значительного числа нефункциональных генов, мРНК которых удаляется компонентами «системы надзора».

У дрожжей деградация требует наличия элементов последовательности (называемых *DSE*), расположенных ниже нонсенс-мутации. Легче всего предположить, что это дестабилизирующие элементы, использование которых подавляется трансляцией. Тем не менее при блокировании трансляции мРНК остается стабильной. Это значит, что процесс деградации напрямую связан с трансляцией мРНК или терминацией синтеза.

Гены, ответственные за этот процесс, были определены у *S. cerivisiae* (локус *upf*) и у *C. elegans* (локус *smg*) путем выявления супрессоров деградации, опосредованной нонсенс-мутациями. Мутации в этих генах стабилизируют аберрантные мРНК, однако не влияют на стабильность большинства транскриптов дикого типа. Гены кодируют АТФ-зависимую хеликазу (фермент, расплетающий двухцепочечные нуклеиновые кислоты). Это значит, что распознавание мРНК, как подходящей мишени для деградации требует изменения ее структуры.

Upf1 взаимодействует с факторами (eRF1 и eRF3), катализирующими терминацию, и, вероятно, распознает стоп-кодоны. Продвигаясь вдоль мРНК к 3′-концу, Upf1 сканирует последовательность молекулы в этом направлении.

В клетках млекопитающих «система надзора» работает только с мутациями, расположенными перед последним экзоном, — иными словами, после мутации должен быть интрон. Это значит, что для работы системы в ядре (до удаления интронов) должно произойти некое событие. Вероятно, белки прикрепляются к мРНК на границе экзонов во время сплайсинга. На РИС. 7.27 показана модель работы такой системы. Она напоминает последовательность событий, в ходе которых происходит мечение мРНК, необходимое для ее экспорта из ядра (см. разд. 26.10 «Сплайсинг связан с экспортом мРНК»). Прикрепление белка к месту соединения экзонов является сигналом к началу событий, которые затем продолжатся в цитоплазме. У человека белки, гомологичные Upf2,3, могут быть вовлечены в работу похожей системы. Они специфично связываются со сплайсированной мРНК.

7.15 РНК эукариот транспортируется через мембрану

Ключевые понятия:

- Трансмембранный транспорт РНК осуществляется в составе рибонуклеопротеиновой частицы.
- Все РНК эукариот, работающие в цитоплазме, должны быть экспортированы из ядра.
- тРНК и РНК-компонент рибонуклеазы импортируются в митохондрию.
- мРНК растений может осуществлять длинные путешествия между клетками.

Бактериальная клетка представляет собой единый компартмент. Таким образом, все функции РНК, а также ее синтез осуществляются в одних тех же условиях. Это особенно удивительно в случае с мРНК, когда трансляция и транскрипция происходят одновременно (см. разд. 7.7 «Жизненный цикл бактериальной мРНК»).

Общая схема транспорта мРНК через мембрану приведена на РИС. 7.28. Для осуществления такого переноса необходимо преодолеть определенное термодинамическое препятствие, возникающее при попытке переноса отрицательно заряженной молекулы РНК через гидрофобную мембрану. Эта проблема решается путем упаковки РНК в белки.

У эукариот РНК транскрибируется в ядре, а трансляция происходит в цитоплазме. Каждый тип РНК должен быть транспортирован в цитоплазму для сборки аппарата трансляции. рРНК объединяются с рибосомальными белками и формируют незрелые субъединицы рибосомы, выполняющие функции субстратов для системы транспорта. тРНК транспортируются посредством специальной белковой системы. мРНК транспортируется в виде рибонуклеопротеина, который формируется в ядре (см. гл. 26 «Сплайсинг и процессинг РНК»). Все эти механизмы являются общими для эукариотических клеток. Многие молекулы мРНК транслируются в цитозоле, однако некоторые из них крепятся к элементам цитоскелета. Одним из примеров такой локализации является синтез белка в непосредственной близости от места его встраивания в некую макромолекулярную структуру.

Некоторые РНК, производимые в ядре, экспортируются в цитозоль, а затем импортируются в митохондрию. Митохондрии некоторых организмов кодируют не все тРНК, необходимые им для синтеза (см. разд. 4.10 «Геномы органелл — кольцевые ДНК, в которых закодирована информация о некоторых белках органелл»). В этих случаях из цитозоля должны быть импортированы

Эукариотическая РНК может перемещаться между клеточными компартментами

РНК	Перемещение	Локализация
Все РНК	Из ядра в цитоплазму	Все клетки
тРНК	Из ядра в митохондрию	Многие клетки
мРНК	Из питающей клетки в ооцит	Эмбриогенез *Drosophila*
мРНК	Между полюсами ооцита	Эмбриогенез *Drosophila*
мРНК	Из клетки в клетку	Флоэма растений

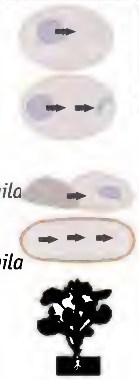


РИС. 7.28 Транспорт РНК через мембрану имеет место в различных системах

недостающие тРНК. Фермент рибонуклеаза Р, содержащий субъединицы белка и РНК, кодируется ядерными генами, однако встречается и в ядре, и в митохондриях. Это значит, что РНК должна импортироваться в митохондрию.

Известны даже случаи межклеточного транспорта мРНК. Во время развития яйцеклетки *Drosophila*, определенные мРНК транспортируются в яйцо из окружающих его питающих клеток. Последние имеют с ооцитом специализированные соединения, которые используют для передачи материала, необходимого для раннего развития. К этому материалу относятся и определенные мРНК. Оказавшись в яйце, эти мРНК распределяются по определенным местам. Некоторые просто располагаются на том конце клетки, где они проникли в нее. Другие же с помощью прикрепленного к микротрубочкам «мотора» транспортируются к противоположному концу яйца.

Самый удивительный случай транспорта мРНК был обнаружен у растений. Перемещение отдельных нуклеиновых кислот на значительные расстояния было впервые описано у растений. Белки некоторых вирусов способствуют распространению инфекции в тканях растений путем переноса вирусной РНК через плазмодесмы (соединения между клетками). Растения же имеют систему защиты, которая позволяет клеткам оставаться невосприимчивыми к инфекции. Похожие системы участвуют и в транспорте мРНК между клетками растений. Несмотря на то что о существовании таких систем известно довольно давно, их функциональная значимость была продемонстрирована сравнительно недавно путем прививания томатов, несущих доминантную мутацию *Me* (вызывающую изменение формы листьев), томатам дикого типа. В итоге мутантная мРНК транспортировалась в листья растений дикого типа и изменяла их форму.

7.16 Местонахождение мРНК может быть четко детерминировано

Ключевые понятия:

- мРНК Ash1 дрожжей формирует рибонуклеопротеин, связывающийся с миозиновым мотором.
- Мотор транспортирует ее вдоль актиновых филаментов в дочернюю клетку.
- В дочерней клетке она заякоривается и транслируется. Таким образом, белок синтезируется только в дочерней клетке.

У эукариот мРНК синтезируется в ядре, а транслируется в цитоплазме. Она попадает в цитоплазму в виде рибонуклеопротеиновой частицы, транспортируемой через ядерные поры. Оказавшись в цитоплазме, мРНК может связываться с рибосомами и транслироваться. Цитоплазма довольно плотно «забита» белками. Непонятно, насколько свободно полисома может продвигаться в цитоплазме и как далеко может уходить мРНК от места проникновения в цитоплазму для того, чтобы быть транслированной. Большая часть мРНК находится в случайных местах, однако место трансляции некоторых из них четко определено. За это ответственны несколько механизмов:

- мРНК может быть транспортирована к месту трансляции.
- Она может быть представлена во многих местах клетки и впоследствии может быть подвержена деградации везде, кроме того места, где она должна быть транслирована.
- мРНК может свободно перемещаться до тех пор, пока не будет захвачена в месте трансляции.

Одним из ярких примеров четкой локализации в пределах клетки является белок дрожжей Ash1. Ash1 подавляет экспрессию НО-эндонуклеазы в отпочковывающейся дочерней клетке. Таким образом, НО экспрессируется только в материнской клетке. В результате тип спаривания меняется только в материнской клетке (см. разд. 19.24 «Переключение контролируется регуляцией эксрпессии НО»). Дочерняя клетка не получает НО потому, что вся мРНК Ash1 транспортируется из материнской клетки (где она синтезируется) в отпочковывающуюся дочернюю клетку.

Мутации в любом из 5 генов, называемых *SHE 1-5*, препятствуют специфической локализации и приводят к тому, что мРНК Ash1 равномерно распределяется между материнским и дочерним компартментами. Белки She1, 2 и 3 связывают Ash1 мРНК в рибонуклеопротеиновую частицу, которая транспортирует мРНК в дочернюю клетку. На РИС. 7.29 показаны функции белков. She1p является миозином (ранее определяемым как Myo4), а She3 и She2 являются белками,

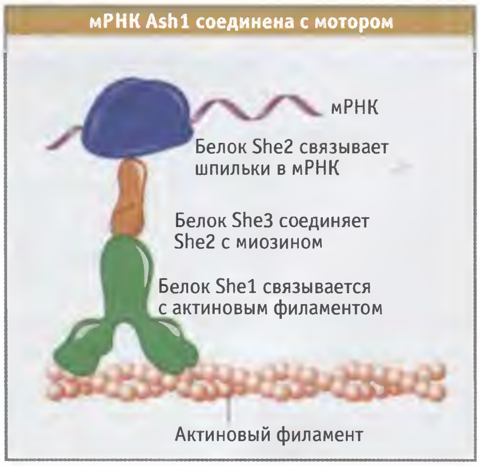


РИС. 7.29 мРНК Ash1 входит в состав нуклеопротеина, который содержит миозиновый мотор, продвигающий ее вдоль актинового филамента



РИС. 7.30 мРНК Ash1 экспортируется из ядра в цитоплазму и там вступает в комплекс с белками She. Комплекс скользит вдоль актиновых филаментов и доставляет мРНК в отпочковывающуюся клетку

соединяющими миозин с мРНК. Миозин выполняет функцию мотора, продвигающего мРНК вдоль актиновых филаментов.

На РИС. 7.30 изображен весь процесс. мРНК Ash1 экспортируется из ядра в виде рибонуклеопротеина. Оказавшись в цитоплазме, она сначала связывается с She2, который распознает некоторые вторичные структуры типа шпилька-петля, расположенные в пределах мРНК. Затем She3 связывается с She2, после чего прикрепляется миозин She1. Далее частица прикрепляется к актиновому филаменту и движется в сторону отпочковывающейся клетки. Достигнув почки мРНК, Ash1 заякоривается там, вероятно, с помощью специфически связанных с молекулой мРНК белков.

Другие примеры доставки мРНК в определенные места в клетке основаны на тех же принципах. мРНК распознается посредством *цис*-действующих последовательностей, которые обычно представлены вторичными структурами, расположенными в 3′-нетранслируемом регионе. (мРНК Ash1 является нетипичным примером, поскольку *цис*-действующий регион здесь расположен в пределах кодирующей рамки.) мРНК упаковывается в рибонуклеопротеиновую частицу (мРНП). Иногда транспортируемую мРНК можно увидеть в очень больших частицах, называемых мРНК гранулами. Эти частицы достаточно велики (в несколько раз больше рибосомы).

Транспортируемый мРНП должен быть соединен с мотором, обеспечивающим его движение по системе «дорожек». Дорожки могут быть представлены актиновыми филаментами или микротрубочками. В то время как Ash1 при движении по актиновым дорожкам использует миозиновый мотор, мРНК *oscar* в яйцеклетке *Drosophila* использует кинезиновый мотор для продвижения по микротрубочкам. Как только мРНК достигает места назначения, ее необходимо заякорить. Этот процесс изучен в меньшей степени, но можно сказать, что он независим от транспорта. мРНК, транспортируемая вдоль микротрубочек, может заякориться на актиновом филаменте в пункте назначения.

7.17 Резюме

Генетическая информация, записанная в ДНК, экспрессируется в две стадии: транскрипция ДНК в мРНК и трансляция мРНК в белок. мРНК транскрибируется с одной из цепей ДНК, она комплементарна этой (некодирующей) цепи и идентична другой (кодирующей) цепи. Последовательность мРНК в направлении 5′–3′ состоит из кодонов по три нуклеотида и соотносится с аминокислотной последовательностью в направлении от N- к С-концу.

Адаптором, определяющим значение кодона, является транспортная РНК, представленная компактной L-образной пространственной структурой; один из концов тРНК несет комплементарный кодону антикодон, а другой ковалентно связан со специфической аминокислотой, соответствующей кодону. тРНК, несущая аминокислоту, называется аминоацил-тРНК.

Рибосома представляет собой аппарат, позволяющий аминоацил-тРНК связываться с соответствующими кодонами мРНК. Малая субъединица рибосомы связана с мРНК; большая субъединица несет формирующийся полипептид. Рибосома движется вдоль мРНК, начиная с сайта инициации, расположенного в 5′-регионе, в сторону сайта терминации в 3′-регионе. По мере ее продвижения соответствующие аминоацил-тРНК связываются с кодонами и отдают свою аминокислоту в строящуюся полипептидную цепь. Таким образом, полипептидная цепь увеличивается на одну аминокислоту с прохождением одного кодона.

Транскрипционный аппарат не является специфичным для ткани или организма; мРНК, взятая из одного источника, может быть транслирована с помощью рибосом и тРНК из другого. Число трансляций данной мРНК является функцией от афинности ее сайта(ов) инициации к рибосомам и от ее стабильности. Существуют примеры, когда трансляция групп мРНК или отдельных мРНК специфически ингибируется: это называют контролем трансляции.

Типичная мРНК включает нетранслируемые 5′-лидер, и 3′-трейлер, и кодирующий регион(ы). Бактериальная мРНК обычно имеет полицистронную структуру с нетранслируемыми регионами, расположенными между цистронами. Каждый цистрон представлен кодирующим регионом, начинающимся специфическим сайтом инициации и заканчивающимся сайтом терминации. Субъединицы рибосомы ассоциируют на участке инициации и диссоциируют на участке терминации каждого из кодирующих регионов.

Растущая бактерия *E. coli* содержит около 20 000 рибосом и 200 000 тРНК (большей частью представленных аминоацил-тРНК). Существует около 1500 молекул мРНК, представляющих собой 2–3 копии каждого из 600 различных мессенджеров.

В трансляции одной мРНК может одновременно принимать участие множество рибосом, формируя полирибосому (или полисому). Крупные бактериальные

полисомы несут десятки рибосом, связанных с одной мРНК. Полисомы эукариот обычно меньших размеров и, как правило, несут менее 10 рибосом; каждая мРНК содержит только одну кодирующую последовательность.

Период полураспада мРНК бактерий обычно мал и занимает всего несколько минут. Трансляция начинается на 5′-конце еще до того, как закончилась транскрипция мРНК. Деградация инициируется эндонуклеазами, отрезающими в 5′–3′ направлении участки последовательности, следующие за рибосомой. После этого экзонуклеазы расщепляют фрагменты до нуклеотидов, разрушая их в направлении от 3′- к 5′-концу. Некоторые последовательности могут стимулировать или предотвращать деградацию бактериальных мРНК.

мРНК эукариот до транспорта в цитоплазму (для трансляции) должна претерпеть процессинг в ядре. К 5′-концу добавляется метилированный кэп. Он состит из нуклеотида, крепящегося к исходному концу 5′–5′ связью, и добавочных метильных групп. Большая часть эукариотической мРНК несет последовательность поли(A), состоящую из ~200 нуклеотидов и расположенную с 3′-конца молекулы. Добавление поли(A) происходит в ядре после транскрипции. Кинетика трансляции и деградации поли(A)$^-$ и поли(A)$^+$ мРНК сходна. мРНК эукариот существует в виде рибонуклеопротеиновой частицы (мРНП); в некоторых случаях мРНП является хранилищем молекул, которые не удалось транслировать. мРНК эукариот обычно сохраняет стабильность на протяжении нескольких часов. Они могут содержать последовательности, запускающие деградацию. Известны случаи регуляции этого процесса.

У дрожжей существует, по крайней мере, два пути деградации мРНК. Оба начинаются с удаления поли(A) с 3′-конца, определяющего потерю поли(A)-связывающего белка, что, в свою очередь, приводит к удалению метилированного кэпа с 5′-конца. По одному из путей мРНК разрушается экзонуклеазами с 5′-конца. По другому пути деградация осуществляется экзосомой — комплексом, включающим несколько экзонуклеаз, и начинается с 3′-конца.

Деградация, опосредованная нонсенс-мутациями, приводит к разрушению мРНК, содержащей нонсенс-кодон перед последним экзоном. Наличие локуса *upf* у дрожжей и *smg* у червей является обязательным условием для запуска процесса. Их продукты обладают хеликазной активностью, необходимой для расплетания мРНК, а также взаимодействия с факторами, останавливающими синтез белка. Особенностью этого процесса у млекопитающих является то, что некоторые из белков прикрепляются к мРНК в ядре по мере удаления интронов в ходе сплайсинга.

В пределах клетки существует механизм транспорта мРНК в определенные места (особенно часто это происходит в ходе эмбрионального развития). В системе Ash1 у дрожжей мРНК транспортируется из материнской клетки в дочернюю посредством миозинового мотора, продвигающего молекулу вдоль актиновых филаментов. У растений мРНК может транспортироваться между клетками.

Литература

7.3 Транспортная РНК по форме напоминает лист клевера

Обзор

Soll, D. and RajBhandary, U. L. (1995). *tRNA Structure, Biosynthesis, and Function.* Washington, DC: American Society for Microbiology.

Исследование

Chapeville, F. et al. (1962). On the role of soluble RNA in coding for amino acids. *Proc. Natl. Acad. Sci. USA* 48, 1086–1092.

Hoagland, M. B. et al. (1958). A soluble RNA intermediate in protein synthesis. *J. Biol. Chem.* 231, 241–257.

Holley, R. W. et al. (1965). Structure of an RNA. *Science* 147, 1462–1465

7.5 мРНК транслируется на рибосомах

Исследование

Dintzis, H. M. (1961). Assembly of the peptide chain of hemoglobin. *Proc. Natl. Acad. Sci. USA* 47, 247–261.

7.6 К одной мРНК присоединяется множество рибосом

Исследование

Slayter, H. S. et al. (1963). The visualization of polyribosome structure. *J. Mol. Biol.* 7, 652–657.

7.7 Жизненный цикл бактериальной мРНК

Исследование

Brenner, S. Jacob, F., and Meselson, M. (1961). An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature* 190, 576–581.

7.9 У эукариот 5′-конец мРНК кэпирован

Обзор

Bannerjee, A. K. (1980). 5′-Terminal cap structure in eukaryotic mRNAs. *Microbiol. Rev.* 44, 175–205.

7.10 3′-конец полиаденилирован

Обзор

Jackson, R. J. and Standart, N. (1990). Do the poly(A) tail and 3′ untranslated region control mRNA translation? *Cell* 62, 15–24.

Исследование

Darnell, J. et al. (1971). Poly(A) sequences: role in conversion of nuclear RNA into mRNA. *Science* 174, 507–510.

7.11 В деградацию мРНК бактерий вовлечено множество ферментов

Обзор

Caponigro, G. and Parker, R. (1996). Mechanisms and control of mRNA turnover in *S. cerevisiae*. *Microbiol. Rev.* 60, 233–249.

Grunberg-Manago, M. (1999). mRNA stability and its role in control of gene expression in bacteria and phages. *Annu. Rev. Genet.* 33, 193–227.

Исследование

Miczak, A., Kaberdin, V. R., Wei, C.-L., and Lin-chao, S. (1996). Proteins associated with RNAase E in a multicomponent ribonucleolytic complex. *Proc. Natl. Acad. Sci. USA* 93, 3865–3869.

O'Hara, E. B. et al. (1995). Polyadenylation helps regulate mRNA decay in *E. coli. Proc. Natl. Acad. Sci. USA* 92, 1807–1811.

Vanzo, N. F. et al. (1998). RNAase E organizes the protein interactions in the *E. coli* RNA degradaosome. *Genes Dev.* 12, 2770–2781.

7.12 Стабильность мРНК зависит от ее структуры и последовательности

Обзор

Coller, J., and Parker, R. (2004). Eukaryotic mRNA decapping. *Annu. Rev. Biochem.* 73, 861–890.

Moore, M. J. (2005). From birth to death: the complex lives of eukaryotic mRNAs. *Science* 309, 1514–1518.

Ross, J. (1995). mRNA stability in mammalian cells. *Microbiol. Rev.* 59, 423–450.

Sachs, A. (1993). Messenger RNA degradation in eukaryotes. *Cell* 74, 413–421.

7.13 В деградацию мРНК вовлечены ферменты с различными активностями

Обзор

Beelman, C. A. and Parker, R. (1995). Degradation of mRNA in eukaryotes. *Cell* 81, 179–183.

Coller, J. and Parker, R. (2004). Eukaryotic mRNA decapping. *Annu. Rev. Biochem.* 73, 861–890.

Jacobson, A. and Peltz, S. W. (1996). Interrelationships of the pathways of mRNA decay and translation in eukaryotic cells. *Annu. Rev. Biochem.* 65, 693–739.

Исследование

Allmang, C., Petfalski, E., Podtelejnikov, A., Mann, M., Tollervey, D., and Mitchell, P. (1999). The yeast exosome and human PM-Scl are related complexes of 3′–5′ exonucleases. *Genes Dev.* 13, 2148–2158.

Bousquet-Antonelli, C., Presutti, C., and Tollervey, D. (2000). Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell* 102, 765–775.

Gao, M., Fritz, D. T., Ford, L. P., and Wilusz, J. (2000). Interaction between a poly(A)-specific ribonuclease and the 5′ cap influences mRNA deadenylation rates *in vitro*. *Mol. Cell* 5, 479–488.

Mitchell, P. et al. (1997). The exosome: a conserved eukaryotic RNA processing complex containing multiple 3′–5′ exoribonuclease activities. *Cell* 91, 457–466.

Muhlrad, D., Decker, C. J., and Parker, R. (1994). Deadenylation of the unstable mRNA encoded by the yeast MFA2 gene leads to decapping followed by 5′–3′ digestion of the transcript. *Genes Dev.* 8, 855–866.

Sheth, U., and Parker, R. (2003). Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science* 300, 805–808.

Tucker, M., Valencia-Sanchez, M. A., Staples, R. R., Chen, J., Denis, C. L., and Parker, R. (2001). The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in *S. cerevisiae. Cell* 104, 377–386.

7.14 Нонсенс-мутации запускают «систему надзора»

Обзор

Hilleren, P. and Parker, R. (1999). Mechanisms of mRNA surveillance in eukaryotes. *Annu. Rev. Genet.* 33, 229–260.

Исследование

Cui, Y., Hagan, K. W., Zhang, S., and Peltz, S. W. (1995). Identification and characterization of genes that are required for the accelerated degradation of mRNAs containing a premature translational termination codon. *Genes Dev.* 9, 423–436.

Czaplinski, K., Ruiz-Echevarria, M. J., Paushkin, S. V., Han, X., Weng, Y., Perlick, H. A., Dietz, H. C., Ter-Avanesyan, M. D., and Peltz, S. W. (1998). The surveillance complex interacts with the translation release factors to enhance termination and degrade aberrant mRNAs. *Genes Dev.* 12, 1665–1677.

Le Hir, H., Moore, M. J., and Maquat, L. E. (2000). Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon–exon junctions. *Genes Dev.* 14, 1098–1108.

Lykke-Andersen, J., Shu, M. D., and Steitz, J. A. (2000). Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. *Cell* 103, 1121–1131.

Peltz, S. W., Brown, A. H., and Jacobson, A. (1993). mRNA destabilization triggered by premature translational termination depends on at least three cis-acting sequence elements and one *trans*-acting factor. *Genes Dev.* 7, 1737–1754.

Pulak, R. and Anderson, P. (1993). mRNA surveillance by the *C. elegans smg* genes. *Genes Dev.* 7, 1885–1897.

Ruiz-Echevarria, M. J. et al. (1998). Identifying the right stop: determining how the surveillance complex recognizes and degrades an aberrant mRNA. *EMBO J.* 15, 2810–2819.

Weng, Y., Czaplinski, K., and Peltz, S. (1996). Genetic and biochemical characterization of mutants in the ATPase and helicase regions of the Upf1 protein. *Mol. Cell Biol.* 16, 5477–5490.

Weng, Y., Czaplinski, K., and Peltz, S. (1996). Identification and characterization of mutations in the *upf1* gene that affect the Upf protein complex, nonsense suppression, but not mRNA turnover. *Mol. Cell Biol.* 16, 5491–5506.

7.15 РНК эукариот транспортируется через мембрану

Обзор

Ghoshroy, S., Lartey, R., Sheng, J., and Citovsky, V. (1997). Transport of proteins and nucleic acids through plasmodesmata. *Ann. Rev. Plant. Physiol. Plant. Mol. Biol.* 48, 27–50.

Jansen, R. P. (2001). mRNA localization: message on the move. *Nat. Rev. Mol. Cell Biol.* 2, 247–256.

Lucas, W. J. and Gilbertson, R. L. (1994). Plasmodesmata in relation to viral movement within leaf tissues. *Ann. Rev. Phytopathol.* 32, 387–411.

Vance, V. and Vaucheret, H. (2001). RNA silencing in plants—defense and counterdefense. *Science* 292, 2277–2280.

Исследование

Kim, M., Canio, W., Kessler, S., and Sinha, N. (2001). Developmental changes due to long-distance movement of a homeobox fusion transcript in tomato. *Science* 293, 287–289.

Puranam, R. S. and Attardi, G. (2001). The RNase P associated with HeLa cell mitochondria contains an essential RNA component identical in sequence to that of the nuclear RNase P. *Mol. Cell Biol.* 21, 548–561.

7.16 Местонахождение мРНК может быть четко детерминировано

Обзор

Chartrand, P., Singer, R. H., and Long, R. M. (2001). RNP localization and transport in yeast. *Annu. Rev. Cell Dev. Biol.* 17, 297–310.

Jansen, R. P. (2001). mRNA localization: message on the move. *Nat. Rev. Mol. Cell Biol.* 2, 247–256.

Kloc, M., Zearfoss, N. R., and Etkin, L. D. (2002). Mechanisms of subcellular mRNA localization. *Cell* 108, 533–544.

Palacios, I. M. and Johnston, D. S. (2001). Getting the message across: the intracellular localization of mRNAs in higher eukaryotes. *Annu. Rev. Cell Dev. Biol.* 17, 569–614.

Исследование

Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S. M., Singer, R. H., and Long, R. M. (1998). Localization of ASH1 mRNA particles in living yeast. *Mol. Cell* 2, 437–445.

Long, R. M., Singer, R. H., Meng, X., Gonzalez, I., Nasmyth, K., and Jansen, R. P. (1997). Mating type switching in yeast controlled by asymmetric localization of ASH1 mRNA. *Science* 277, 383–387.

Синтез белка

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

8.12 При транслокации рибосома приходит в движение

- При транслокации мРНК протягивается сквозь рибосому на три основания.
- При транслокации деацилированная тРНК переходит в сайт E, а пептидил-тРНК — в сайт P; сразу после транслокации сайт A оказывается незанятым.
- Модель гибридного состояния строится на предположении, что транслокация происходит в два этапа: сначала 50S субъединица рибосомы сдвигается относительно 30S субъединицы, затем 30S субъединица сдвигается относительно мРНК, возвращая рибосоме ее первоначальную конформацию.

8.13 Факторы элонгации поочередно связываются с рибосомой

- Для транслокации нужен фактор EF-G, по своей структуре похожий на комплекс EF-Tu–ГТФ–аминоацил-тРНК.
- Рибосома не может одновременно связывать EF-Tu и EF-G.
- Для транслокации используется гидролиз ГТФ, которым укомплектован EF-G: эта реакция, меняющая конформацию EF-G, ускоряет изменение структуры рибосомы.

8.14 Синтез белка терминируется тремя кодонами

- Каждый из трех кодонов UAA (охра), UAG (янтарь) и UGA (опал) может терминировать синтез белка.
- Для большинства бактерий, в порядке убывания частоты использования, терминирующие кодоны можно расположить так: UAA>UGA>UAG.

8.15 Терминирующие кодоны узнаются белковыми факторами

- Терминирующие кодоны узнаются специальными белками (факторами терминации, или релизинг-факторами), а не аминоацил-тРНК.
- Релизинг-факторы класса 1 по структуре напоминают аминоацил-тРНК–EF-Tu и EF-G.
- Релизинг-факторы класса 1 узнают терминирующие кодоны; это узнавание выливается в гидролиз связи между полипептидом и несущей его тРНК.
- Релизинг-факторам класса 1 помогают ГТФ-зависимые релизинг-факторы класса 2.
- У бактерий (обладателей двух разновидностей релизинг-факторов класса 1) и эукариот (с однотипными релизинг-факторами класса 1) механизмы терминации, в целом, похожи.

8.16 Обе субъединицы рибосомы пронизаны рибосомальными РНК

- Каждая молекула рибосомальной РНК (рРНК) имеет несколько дискретных доменов, формирующих независимые складки.
- Практически все рибосомальные белки находятся в контакте с рРНК.
- Большинство контактов между субъединицами рибосомы — это контакты между 16S и 23S рРНК.

8.17 В рибосоме есть несколько активных центров

- Взаимодействия с участием рРНК служат главной движущей силой в работе рибосомы.
- Облик тРНК-связывающих сайтов в значительной степени определяется свойствами конкретных участков рРНК.

8.18 16S рРНК играет активную роль в синтезе белка

- 16S тРНК играет активную роль в функционировании 30S субъединицы. 16S рРНК напрямую взаимодействует с мРНК, 50S субъединицей, а также с антикодонами тРНК в сайтах P и A.

8.19 23S рРНК обладает пептидилтрансферазной активностью

- Активность пептидилтрансферазы, присущая рибосоме, заключена в 23S рРНК.

8.20 Встреча большой и малой субъединиц меняет их структуру

- При образовании полной рибосомы головка 30S субъединицы совершает поворот на шее.
- В полных рибосомах активный центр пептидил-трансферазы в составе 50S субъединицы работает гораздо эффективнее, чем в свободных 50S субъединицах.
- В области взаимодействия 30S и 50S субъединиц есть множество участков, доступных растворителю.

8.21 Резюме

8.1 Введение

Каждая молекула мРНК состоит из серии кодонов, взаимодействие которых с антикодонами аминоацил-тРНК приводит к объединению связанных с ними аминокислот в полипептидную цепь. Рибосома контролирует процесс связывания мРНК с аминоацил-тРНК. Рибосома похожа на небольшую мобильную фабрику, которая шагает вдоль матрицы, образуя пептидные связи. Доставляющие аминокислоты аминоацил-тРНК проскакивают сквозь рибосому с умопомрачительной скоростью по мере того, как с ней циклически ассоциируют и диссоциируют факторы элонгации. Рибосома вместе со вспомогательными факторами обеспечивает полный набор активностей для всех этапов синтеза белка.

На рис. 8.1 показано соотношение физических размеров различных компонентов аппарата синтеза белка. Рибосома состоит из двух субъединиц; каждая субъединица играет в синтезе белка свою, особую роль. Матричная РНК ассоциирована с малой субъединицей рибосомы (в каждый момент синтеза ~30 оснований мРНК взаимодействует с малой субъединицей). Молекула мРНК пролегает по поверхности малой субъединицы в непосредственной близости от области контакта субъединиц. В каждый момент времени в синтезе активно участвуют *две* молекулы тРНК, соответствующие двум смежным кодонам. Эти две тРНК взаимодействуют с рибосомой по внутренним сайтам, образуемым сразу обеими субъединицами. Таким образом, в каждом шаге элонгации полипептида

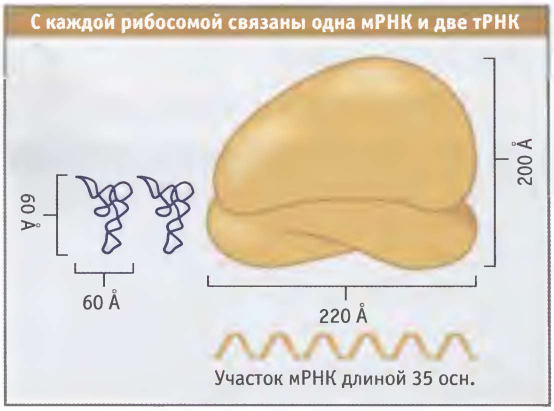


РИС. 8.1 Сравнение размеров показывает, что рибосома достаточно велика для связывания мРНК и молекул тРНК

задействовано всего два кодона из примерно десяти кодонов (~30 осн. мРНК), покрываемых рибосомой. В рибосоме может присутствовать еще и третья тРНК, уже отдавшая свою аминокислоту растущему белку: она «медлит», прежде чем уйти из рибосомы в цитоплазму для «перезарядки».

В ходе эволюции живых организмов общий план строения рибосомы почти не менялся. Тем не менее рибосомы бактерий, рибосомы цитоплазмы эукариот и рибосомы органелл имеют множество различий (по размеру, содержанию рибосомальных белков и т. п.). В таблице на **РИС. 8.2** приведено сравнение рибосом бактерий и млекопитающих. У тех и других (как и у всех промежуточных, с точки зрения эволюции, организмов) рибосомы — это рибонуклеопротеиновые частицы, в которых больше РНК, чем белка. Рибосомальные (или рибосомные) белки иногда называют *r-белки*.

Каждая субъединица рибосомы содержит преобладающую по размерам рРНК и определенное число небольших белков. Большая субъединица рибосомы может также содержать одну или несколько небольших рРНК. Так, у *E. coli* малая субъединица (30S) состоит из одной рРНК (16S) и двадцати одного r-белка. Большая (50S) субъединица рибосомы *E. coli* содержит две рРНК (23S и 5S) и тридцать один r-белок. За исключением одного r-белка (который присутствует в количестве четырех копий), каждый r-белок представлен в рибосоме единственной копией. Основную массу бактериальной рибосомы составляют большие рРНК. Рибосома буквально пронизана ими, и большинство r-белков (возможно, все) контактирует непосредственно с рРНК. Таким образом, большие рРНК образуют остов субъединиц. Особая укладка непрерывных нитей больших рРНК, задавая расположение r-белков, доминирует над всей структурой рибосомы.

Эукариотические рибосомы отличаются более высоким содержанием и белков, и РНК: они крупнее, чем бактериальные. В них насчитывается больше r-белков. Скорее всего, в каждой рибосоме большинство эукариотических r-белков (если не все) представлено несколькими копиями (согласно данным стехиометрии их взаимодействия с рРНК). Большие эукариотические рРНК (28S и 18S) длиннее соответствующих прокариотических рРНК (23S и 16S). Основную массовую долю эукариотической рибосомы по-прежнему составляет РНК.

В органеллах эукариотической клетки также есть собственные рибосомы. Они отличаются от рибосом цитоплазмы и могут быть нескольких видов. В некоторых случаях они содержат 70% РНК и напоминают бактериальные рибосомы; иногда они еще мельче (60S) и содержат менее 30% РНК.

В каждой рибосоме есть несколько активных центров. Каждый активный центр образован группой r-белков, находящихся в контакте с конкретным участком рРНК. В работе активных центров всегда участвует рРНК; ее роль может быть не только структурообразующей, но и каталитической. Разумеется, каталитические функции рибосомы требуют участия специальных белков; тем не менее ни одну из активностей, присущих рибосоме, нельзя воспроизвести *in vitro* силами одних только белков (даже объединяя их в группы). Рибосомальные белки функционируют только в контексте рибосомы.

Наши знания о рибосоме базируются на двух типах данных: результатах анализа мутаций и анализа структуры рибосомы. Мутации указывают на определенные r-белки или на определенные участки в рРНК как на участников конкретных реакций. Структурный анализ, в том числе с использованием прямой химической модификации компонентов рибосомы и всевозможных сравнений, направлен на идентификацию функциональных и консервативных особенностей рРНК. Эти исследования помогают представить и уточнить пространственную организацию рибосомы как сложнейшей системы компонентов, задействованных в выполнении конкретных функций.

Рибосомы – это рибонуклеопротеиновые частицы

Рибосомы		рРНК	Рибосомальные белки
Бактериальная (70S) масса 2,5 МДа, 66% рРНК	50S	23S = 2904 осн. 5S = 120 осн.	31
	30S	16S = 1542 осн.	21
Млекопитающих (80S) масса 4,2 МДа, 60% рРНК	60S	28S = 4718 осн. 5,8S = 160 осн. 5S = 120 осн.	49
	40S	18S = 1874 осн.	33

РИС. 8.2 Рибосомы — это огромные рибонуклеопротеиновые комплексы, в которых больше РНК, чем белка. Рибосомы диссоциируют на большую и малую субъединицы

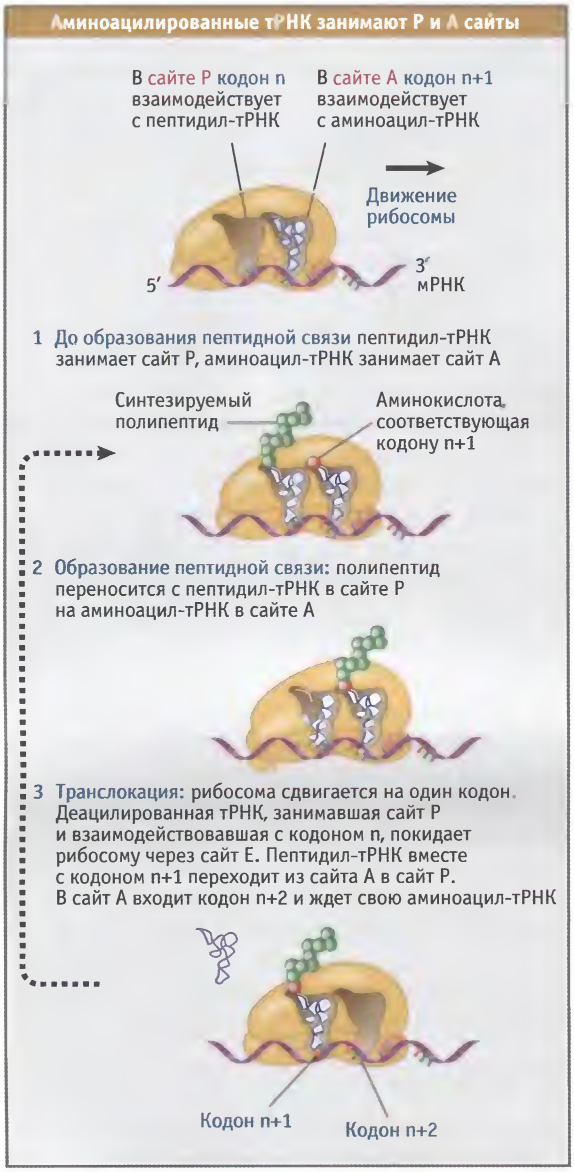


РИС. 8.3 В рибосоме есть два сайта для связывания «заряженной» тРНК



РИС. 8.4 Сайты Р и А устроены таким образом, что взаимодействующие тРНК располагаются поперек обеих субъединиц рибосомы

8.2 Синтез белка состоит из инициации, элонгации и терминации

Ключевые понятия:

- На рибосоме есть три сайта связывания тРНК.
- Аминоацил-тРНК садится на сайт А.
- Пептидил-тРНК находится в сайте Р.
- Деацилированная тРНК выходит через сайт Е.
- Добавление очередной аминокислоты к растущему полипептиду осуществляется за счет его переноса с пептидил-тРНК в сайте Р на аминоацил-тРНК в сайте А.

При синтезе белка мономерами-предшественниками полипептида служат аминокислоты. Они попадают в рибосому не сами по себе, а в комплексе с тРНК. Добавление новой аминокислоты к растущей полипептидной цепочке происходит в результате реакции между тРНК, несущей эту аминокислоту, и тРНК, принесшей предыдущую аминокислоту. Каждая из этих двух тРНК занимает в рибосоме свой отдельный сайт, как показано на РИС. 8.3.

- Входящая аминоацил-тРНК связывает **сайт А**, где ее уже поджидает следующий по очереди кодон мРНК. Именно этот кодон (в сайте А) диктует рибосоме тип входящей аминоацил-тРНК (т. е. следующую аминокислоту растущего полипептида).
- Предыдущий кодон соответствует аминокислоте, которая только что была добавлена к насцентному (растущему) полипептиду. Этот кодон расположен в **сайте Р** вместе с **пептидил-тРНК**; все еще взаимодействуя с этим кодоном, пептидил-тРНК несет на себе всю насцентную полипептидную цепь.

На РИС. 8.4 хорошо видно, что аминоацильные концы двух тРНК находятся в большой субъединице рибосомы, в то время как их антикодоны взаимодействуют с мРНК на поверхности малой субъединицы. Таким образом, каждый из сайтов Р и А образован совместно большой и малой субъединицами рибосомы.

Чтобы рибосома могла приступить к образованию новой пептидной связи, ей нужно прийти в состояние, изображенное в верхней части рис. 8.3, где пептидил-тРНК занимает сайт Р, а аминоацил-тРНК занимает сайт А. Образование пептидной связи происходит в тот момент, когда растущий полипептид переносится со своей (пептидил-) тРНК на аминокислоту аминоацил-тРНК (из сайта Р в сайт А). Эту реакцию катализирует большая субъединица рибосомы.

Перенос пептида на тРНК, находящуюся в сайте А, меняет состояние всей рибосомы (рис. 8.3). **Деацилированная тРНК** (бывшая пептидил-тРНК, теперь лишившаяся всех аминокислот) продолжает занимать сайт Р. В сайте А теперь образовалась новая пептидил-тРНК (с новой пептидной связью между первыми двумя аминокислотными остатками, считая от тРНК, и длиннее на одно звено полипептидной цепочки).

После этого рибосома сдвигается вдоль мРНК на один триплет (один кодон). Данную ступень рабочего цикла

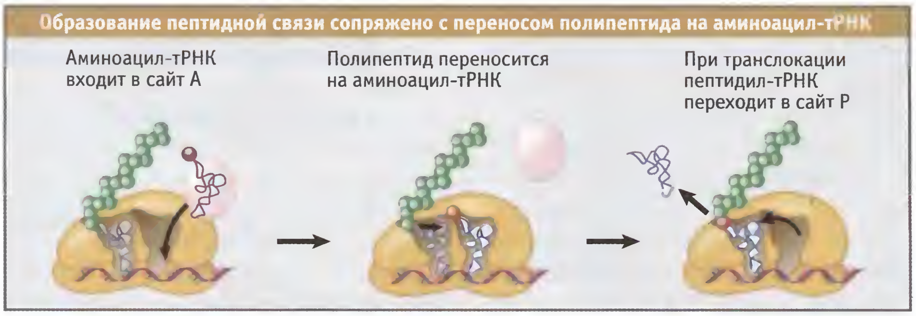


РИС. 8.5 В каждом рабочем цикле элонгации аминоацил-тРНК входит в сайт А, принимает на себя полипептид и переходит в сайт Р

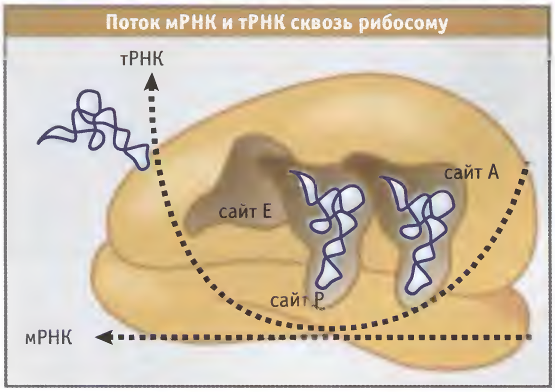


РИС. 8.6 тРНК и мРНК проходят сквозь рибосому в одном и том же направлении

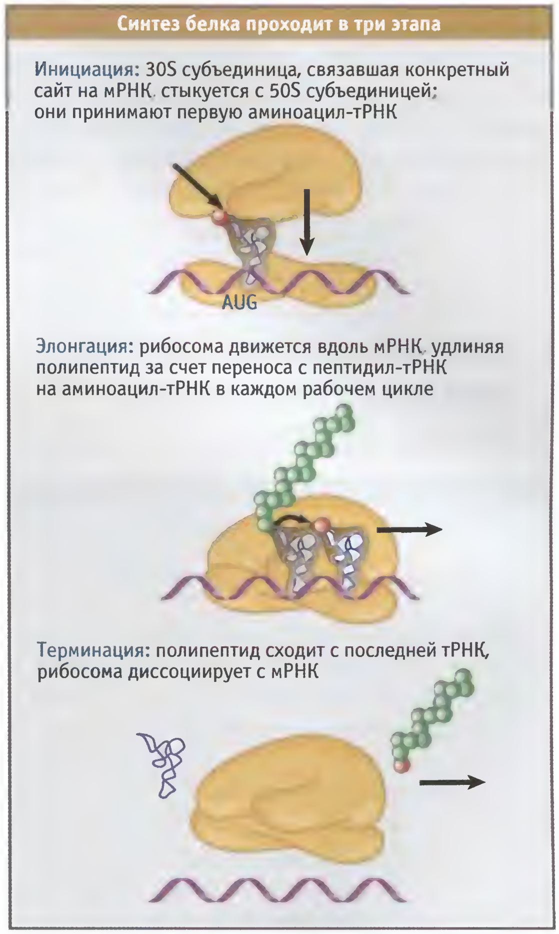


РИС. 8.7 Синтез белка проходит в три этапа

рибосомы называют транслокацией. При транслокации деацилированная тРНК уходит из сайта Р; на ее место переезжает пептидил тРНК из сайта А (см. рис. 8.3). Теперь в сайте А находится следующий, пока не использованный кодон, готовый принять свою аминоацил-тРНК; он будет транслирован в следующем рабочем цикле рибосомы.

Все взаимодействия между тРНК и рибосомой изображены на РИС. 8.5. Отметим, что деацилированная тРНК покидает рибосому через еще один сайт связывания, Е. Этот сайт удерживает тРНК недолго. Таким оборазом, существует как бы поток тРНК сквозь рибосому: молекула тРНК входит в сайт А, из него переходит в сайт Р и затем, через сайт Е, выходит наружу (см. также рис. 8.28 в разд. 8.12). На РИС. 8.6 показано движение пула тРНК относительно молекулы мРНК; оно похоже на зубчатую передачу (если считать взаимодействие кодона с антикодоном эквивалентом механического зацепления).

Весь процесс синтеза одного полипептида состоит из трех этапов, изображенных на РИС. 8.7.

- **Инициация** — это система реакций, ведущих к образованию пептидной связи между первыми двумя аминокислотными остатками нового полипептида. Задача, на самом деле, очень сложная: чтобы первая аминоацил-тРНК могла вступить в комплекс инициации, нужно, чтобы рибосома сначала связалась с мРНК. Инициация — сравнительно медленный этап синтеза белка, и обычно именно инициация задает скорость трансляции мРНК.
- **Элонгация** — это все, что происходит с растущим полипептидом за время образования всех пептидных связей (от связи между первыми двумя аминокислотными остатками до момента добавления последней аминокислоты). На каждом шаге элонгации

рибосома, совершая один рабочий цикл, добавляет к растущему полипептиду одно новое звено (что происходит очень быстро — это, пожалуй, самая быстрая стадия во всем процессе синтеза белка).

- **Терминация** включает все реакции, направленные на освобождение синтезированного полипептида и диссоциацию рибосомы с мРНК.

Каждому этапу синтеза белка соответствует особый набор вспомогательных факторов, обслуживающих рибосому. Источником энергии на всех этапах служит гидролиз ГТФ.

При инициации малая субъединица рибосомы связывает мРНК и только после этого соединяется с большой субъединицей. При элонгации мРНК протягивается сквозь рибосому, которая транслирует ее триплет за триплетом. (Иногда бывает удобно считать, что транслируемая мРНК неподвижна, а движется рибосома; но в этой главе будет уместнее выбрать систему отсчета, связанную с рибосомой.) При терминации готовый полипептид высвобождается, мРНК уходит, рибосома (может быть, не сразу) диссоциирует на отдельные субъединицы, которые еще много раз будут использованы при синтезе белка.

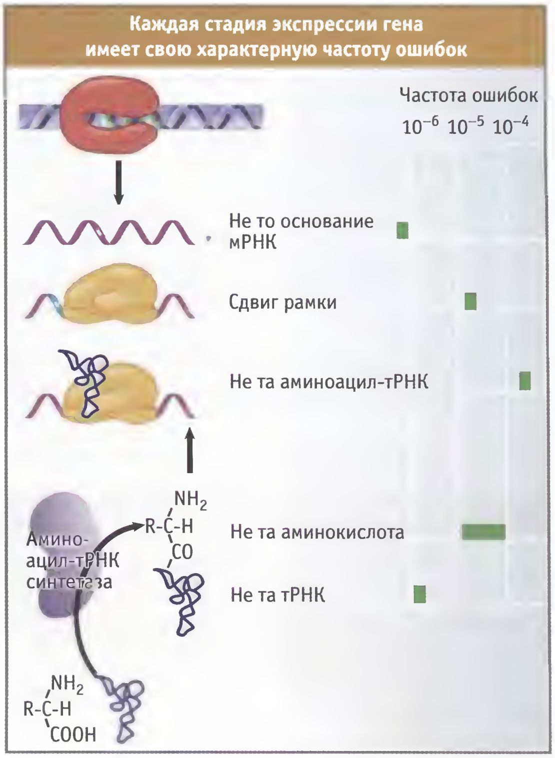


РИС. 8.8 Ошибки синтеза возникают с частотой от 10^{-6} до 5×10^{-4} на разных стадиях экспрессии

8.3 Специальные механизмы контролируют точность синтеза белка

Ключевые понятия:

- Для каждого этапа синтеза белка существуют специфические механизмы, контролирующие его точность.

Легко убедиться, что синтез белка — очень точный процесс (потому что любая молекула конкретного белка имеет одну и ту же последовательность аминокислот). На самом деле ошибки случаются; согласно (немногочисленным) тщательным измерениям *in vivo*, ошибочная аминокислота встраивается в белок не чаще, чем один раз на 10^4, но и не реже, чем один раз на 10^5. Ошибки не настолько многочисленны, чтобы влиять на фенотип (особенно если учесть, что белки продуцируются клеткой в значительных количествах, подавляющее большинство которых безупречны).

Ошибок действительно *очень мало*; до сих пор не совсем понятно, за счет чего обеспечивается такая аккуратность? Почему аминоацил-синтетазы распознают тРНК и аминокислоты с такой огромной точностью? Каким образом рибосома так безошибочно отличает тРНК, соответствующую кодону в сайте А? Можно посмотреть еще шире: в живой клетке последовательность звеньев многих полимеров несет уникальную смысловую нагрузку. Именно *точность* матричного синтеза обеспечивает выживание. Репликация, транскрипция, трансляция... В каждом случае решается одна и та же задача: выбор единственно верного мономера-предшественника из набора других, *похожих*, мономеров. Причем в хорошем темпе и бесчисленное множество раз подряд — для *каждого положения* в полимере.

Скорее всего, общий принцип всегда один и тот же: в начале каждого шага синтеза в контакт с активным центром фермента позволено вступать всем мономерам. Но остаться в активном центре синтеза может только подходящий мономер. Все остальные мономеры, так или иначе, получают отказ в длительном контакте с активным центром. Подходящие мономеры всегда составляют меньшинство (одна из двадцати аминокислот, одна из сорока тРНК, один из четырех нуклеотидов), поэтому критерии дискриминации должны быть очень строгими. Фермент должен уметь каким-то образом амплифицировать различия между потенциальными субстратами, поскольку нет уровня точности, который может быть достигнут установлением пробных контактов с доступными поверхностями потенциального субстрата.

На **РИС. 8.8** приведено сравнение частот возникновения ошибок на разных этапах перевода генетической информации в аминокислотную последовательность белка.

При транскрипции ошибки происходят с частотой $< 10^{-6}$. Контроль качества копирования на этом этапе экспрессии особенно важен, так как *одна* молекула мРНК транслируется множество раз (множеством рибосом), отображаясь во *множество* идентичных полипептидов. Не совсем понятно, каким образом РНК-полимеразе удается настолько редко ошибаться.

Рибосома, в свою очередь, ошибается чуть чаще. Она может делать ошибки двух типов. Во-первых, пропуск одного основания мРНК вызывает сдвиг рамки считывания (frameshift). Сдвиг рамки считывания в направлении, обратном ходу трансляции, иногда случается, когда рибосома читает одно основание дважды: как последнее в кодоне и как первое в следующем кодоне. В норме такие ошибки случаются с частотой $\sim 10^{-5}$. Второй тип ошибок, совершаемых рибосомой, заключается в несоответствии аминоацил-тРНК кодону мРНК. Результат такой ошибки — замена одного аминокислотного остатка другим. Это самая распространенная ошибка синтеза белка, ее частота равна $\sim 5 \times 10^{-4}$. Держать ситуацию под контролем позволяют структура рибосомы и скорость трансляции (см. разд. 9.15 «Рибосомы влияют на точность транскрипции»).

Фермент аминоацил-тРНК-синтетаза тоже может ошибаться. По идее, этот фермент специфичен и к аминокислоте, и к тРНК. Так что ошибки бывают двух типов: либо фермент берет не ту аминокислоту, либо он берет не ту тРНК. Поскольку у тРНК больше пробных контактов с аминоацил-тРНК-синтетазой (фактически она предоставляет ферменту целую поверхность для сканирования), фермент путает тРНК реже, чем аминокислоты. Аминоацил-тРНК-синтетаза располагает дополнительными механизмами исправления ошибок *до* того, как тРНК, заряженная не той аминокислотой, отправится участвовать в синтезе белка (см. разд. 9.11 «Синтетазы — ферменты с корректирующей активностью»).

Свободная рибосома может диссоциировать на отдельные субъединицы, большую и малую. Таким образом, 70S рибосомы находятся в динамическом равновесии с 50S и 30S субъединицами. *Инициация синтеза белка не является функцией интактных рибосом: большая и малая субъединицы рибосомы участвуют в инициации по отдельности.* Реассоциация субъединиц происходит в течение инициации. На **РИС. 8.9** показан цикл рибосомальных субъединиц, участвующих в синтезе белка у бактерий.

Инициация происходит на особом участке мРНК со специфической последовательностью. Этот небольшой участок известен как **сайт связывания рибосомы**; он находится непосредственно перед кодирующим регионом (см. рис. 8.1). Сайт связывания рибосомы служит местом встречи большой и малой субъединиц будущей интактной рибосомы. Реакция соединения субъединиц проходит в два этапа:

- Первым этапом можно считать узнавание мРНК, выражающееся в связывании малой субъединицы с мРНК, с образованием **комплекса инициации**.
- К комплексу инициации присоединяется большая субъединица, формируя целую рибосому.

30S субъединица, вовлеченная в инициацию, сама по себе не способна обеспечить реакции связывания мРНК и тРНК. Для этого привлекаются дополнительные белки, называемые **факторами инициации** (IF). Их можно обнаружить в комплексе с 30S субъединицей, но не с целой 70S рибосомой: когда субъединицы соединяются, факторы инициации

8.4 В инициации синтеза белка у бактерий участвуют 30S субъединицы и вспомогательные факторы

Ключевые понятия:

- Для инициации 30S и 50S субъединицы рибосомы должны находиться отдельно друг от друга.
- Также необходимы факторы инициации (IF-1, IF-2 и IF-3), которые связываются с 30S субъединицей.
- 30S субъединица, несущая факторы инициации, садится на сайт инициации мРНК; так образуется комплекс инициации.
- Чтобы 50S субъединица могла приосединиться к комплексу 30S-мРНК, IF-3 должен уйти.

Рибосомы бактерий, осуществляющие элонгацию полипептида, представляют собой частицы с коэффициентом седиментации 70S. При терминации эти 70S частицы сходят с мРНК в интактном виде как свободные рибосомы. В растущих бактериях большинство рибосом активно участвует в синтезе белка; пул свободных рибосом составляет всего ~20% от общего пула рибосом.

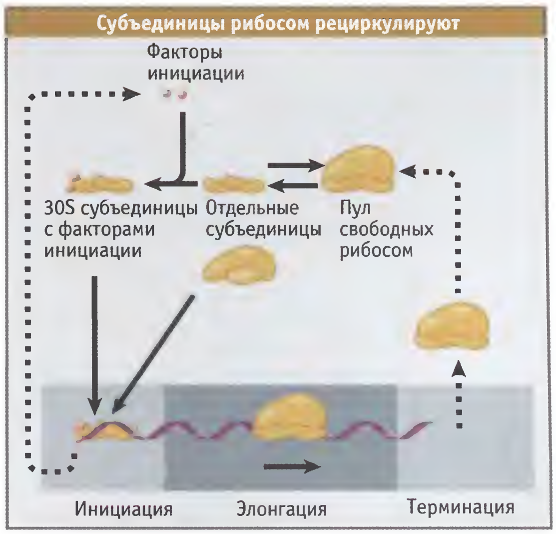


РИС. 8.9 Для инициации требуются свободные субъединицы рибосом. Когда рибосома высвобождается при терминации, 30S субъединицы улавливаются факторами инициации и диссоциируют с образованием свободных субъединиц. Когда при инициации трансляции субъединицы реассоциируют с образованием функциональной рибосомы, факторы инициации уходят

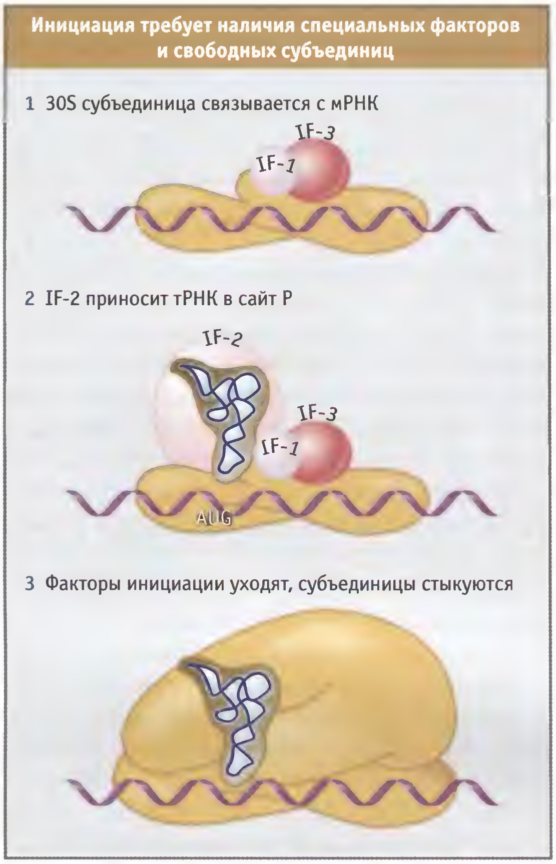


РИС. 8.10 Факторы инициации стабилизируют свободные 30S субъединицы и ставят инициаторную тРНК на ее место в комплексе 30S-мРНК

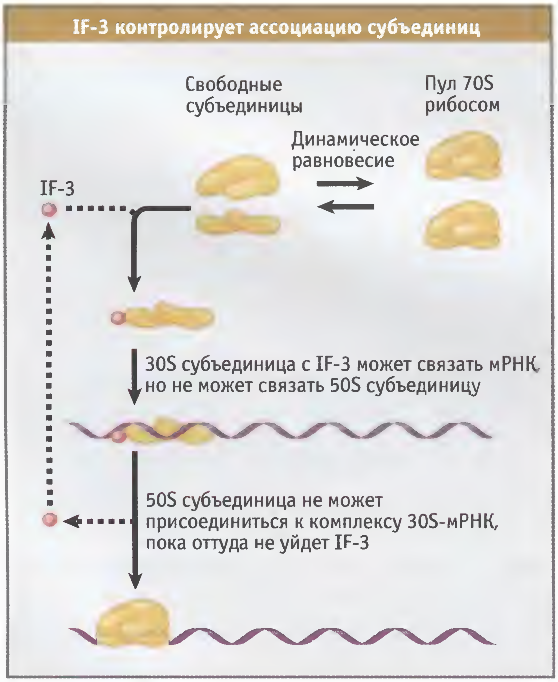


РИС. 8.11 Для инициации требуются 30S субъединицы, несущие IF-3

уходят. В этом их отличие от рибосомальных белков (r-белков, структурных белков рибосомы). Факторы инициации нужны только для образования комплекса инициации; их нет в 70S рибосомах; они не участвуют в элонгации.

Этап инициации полностью изображен на РИС. 8.10; его можно разделить на три шага. У бактерий есть три фактора инициации; им присвоены номера **IF-1**, **IF-2** и **IF-3**. Без этих факторов ни мРНК, ни тРНК не могут вступить в комплекс инициации.

- IF-3 нужен для того, чтобы 30S субъединицы могли связывать специфические сайты мРНК.
- IF-2 связывает особую инициаторную тРНК и контролирует ее вхождение на рибосому.
- IF-1 связывается с 30S субъединицами в уже сформированном комплексе 30S-мРНК. IF-1 связывается с сайтом А, не позволяя аминоацил-тРНК занять этот сайт, и, возможно, отодвигает момент связывания инициаторного комплекса с 50S субъединицей.

У IF-3 множество функций. Вначале он стабилизирует свободные 30S субъединицы. Затем придает им способность связаться с мРНК. Наконец, уже в составе комплекса 30S-мРНК, IF-3 проверяет правильность соответствия первой аминоацил-тРНК старт-кодону (см. разд. 8.6 «Использование fMet-тРНК$_f$ контролируется фактором IF-2 и рибосомой»).

Как показано на РИС. 8.11, первая функция IF-3 — это смещение равновесия между состояниями рибосом. IF-3 связывает свободные 30S субъединицы на выходе из пула 70S рибосом (в результате диссоциации последних). Присутствие IF-3 мешает воссоединению 30S субъединицы с 50S субъединицей. С каждой 30S субъединицей связывается одна молекула IF-3. Число свободных 30S субъединиц определяется доступностью IF-3, которого в клетке всегда сравнительно мало.

IF-3 связывается с поверхностью 30S субъединицы неподалеку от сайта А. Существует значительное совпадение между основаниями 16S рРНК, покрытыми IF-3, и теми основаниями, которые задействованы при связывании с 50S субъединицей. Похоже на то, что IF-3 физически мешает связыванию большой и малой субъединиц рибосомы. Таким образом, IF-3 работает как фактор, противодействующий ассоциации субъединиц и помогающий 30S субъединице сохранять свободную форму.

Вторая функция IF-3 касается способности 30S субъединиц к связыванию с мРНК. IF-3 нужен малым субъединицам для образования комплекса с мРНК. Чтобы к этому комплексу смогла присоединиться 50S субъединица, IF-3 должен уйти. Как только IF-3 уходит, он сразу же взаимодействует с новой свободной 30S субъединицей.

Другой фактор инициации, IF-2, обладает ГТФазной активностью (но проявляет эту активность только в комплексе с рибосомой), то есть, в присутствии рибосомы IF-2 проводит расщепление (гидролиз) ГТФ

с высвобождением энергии. ГТФ гидролизуется в тот момент, когда 50S субъединица присоединяется к комплексу инициации с образованием полной рибосомы. Расщепление ГТФ ведет к изменению конформации субъединиц, способствуя их превращению в активную 70S рибосому.

8.5 Особая инициаторная тРНК закладывает первое звено полипептида

Ключевые понятия:

- Каждый синтезируемый полипептид начинается с аминокислоты метионина, обычно кодируемой триплетом AUG.
- В инициации и элонгации участвуют разные метиониновые тРНК.
- У инициаторной тРНК есть структурные особенности, которыми она отличается от всех остальных тРНК.
- У бактерий метионин инициаторной тРНК имеет формилированную аминогруппу.

Синтез всех белков начинается с одной и той же аминокислоты — метионина. Сигналом к началу синтеза полипептидной цепи служит особый инициирующий кодон мРНК, обозначающий начало открытой рамки считывания. Обычно таким кодоном служит триплет AUG; у бактерий, помимо AUG, используются триплеты GUG и UUG.

Хотя в открытой рамке считывания кодоны AUG всегда соответствуют метионинам, для них существует два принципиально разных вида тРНК. Один вид тРНК используется для инициации, другой — для узнавания внутренних кодонов AUG во время элонгации.

У бактерий, а также в органеллах эукариот инициаторная тРНК несет метиониновый остаток, который формилирован по аминогруппе. Таким образом, вся молекула есть не что иное, как **N-формил-метионил-тРНК**. Сама по себе инициаторная тРНК обозначается как **тРНК$_f^{Met}$**. Аминоацилированная инициаторная тРНК обозначается как **fMet-тРНК$_f$**.

Инициаторная тРНК получает свою формилированную аминокислоту в результате реакции, проходящей в две ступени. Сначала тРНК «заряжается» обычной аминокислотой с образованием Met-тРНК$_f$. Затем идет реакция формилирования, блокирующая свободную аминогруппу метионила, как показано на РИС. 8.12. Полученная fMet-тРНК$_f$ лишена возможности участвовать в элонгации (из-за блока аминогруппы), зато хорошо приспособлена к участию в закладке будущего полипептида.

Такая тРНК используется только в инициации. Она узнает кодоны AUG и GUG (иногда UUG). Качество узнавания зависит от кодона: при замене AUG на GUG эффективность инициации падает вдвое, при замене GUG на UUG — еще вдвое.

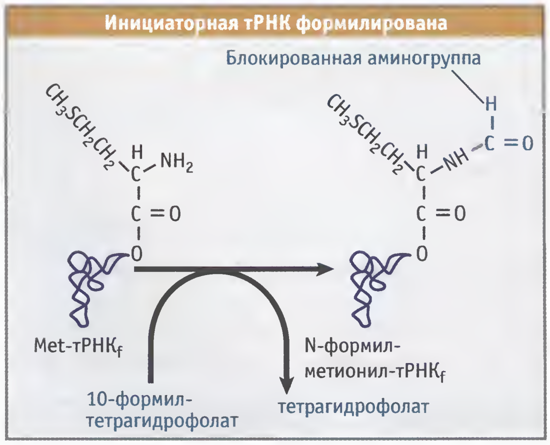


РИС. 8.12 Инициаторная N-формил-метионил-тРНК (fMet-tRNA$_f$) образуется за счет формилирования метионил-тРНК. В качестве кофактора в этой реакции используется формилтетрагидрофолат



РИС. 8.13 fMet-тРНК$_f$ обладает рядом уникальных особенностей, отличающих ее как инициаторную тРНК

За узнавание кодона AUG во внутренних положениях открытой рамки считывания отвечает другой вид **тРНК**. Антикодон этой тРНК взаимодействует только с внутренними AUG; метионил, который она несет, не может быть формилирован.

В чем отличие инициаторной fMet-тРНК$_f$ от элонгаторной Met-тРНК$_m$? Как показано на РИС. 8.13, их нуклеотидные последовательности имеют ряд характерных особенностей. Так, некоторые особенности последовательности инициаторной тРНК мешают учас-

твовать в элонгации, другие, напротив, необходимы для участия в инициации:

- Формилирование не является строго необходимым, поскольку неформилированная Met-тРНК$_f$ тоже может работать как инициатор. Но формилирование значительно повышает эффективность использования Met-тРНК$_f$, поскольку формил служит одним из сигналов для связывания IF-2 с инициаторной тРНК.
- Все тРНК, кроме тРНК$_f^{Met}$, имеют исключительно спаренные основания в акцепторном участке тРНК, где она соединяется с аминокислотой. Мутации, придающие этой паре комплементарность, позволяют тРНК$_f^{Met}$ участвовать в элонгации. Следовательно, присутствие неспаренных оснований именно в этом положении фактически удерживает тРНК$_f^{Met}$ от участия в элонгации. Установлено, что эта пара также мешала бы реакции формилирования метионила.
- Еще одной отличительной особенностью тРНК$_f^{Met}$ является серия из трех комплементарных G-C в шпильке антикодоновой петли. Эти пары G-C необходимы для посадки fMet-тРНК$_f$ в сайт P на малой субъединице рибосомы.

В бактериях и митохондриях есть специальный фермент, деформилаза, «обкусывающий» формильную группу первого метионина, так что у растущего полипептида получается нормальный NH_2-конец. В половине случаев сам метионин также подлежит удалению: фермент аминопептидаза удаляет у полипептида деформилированный N-концевой метионин; таким образом, вторая аминокислота становится первой. Деформилирование и, если положено, удаление деформилированного N-концевого метионина происходят вскоре после инициации, когда длина растущего полипептида еще не достигла и 15 аминокислот.

8.6 Использование fMet-тРНК$_f$ контролируется фактором IF-2 и рибосомой

Ключевые понятия:

- IF-2 связывает инициаторную fMet-тРНК и позволяет ей занять неполный сайт P на малой субъединице рибосомы.

Трансляция кодонов AUG зависит от **контекста**. Если AUG используется для инициации, он читается как формил-метионин; если AUG стоит внутри открытой рамки считывания, он обозначает метионин. Для трансляции кодона GUG окружение еще более критично. В качестве первого кодона GUG транслируется в формил-метионин; в качестве внутреннего кодона GUG читается Val-тРНК и обозначает, естественно, валин (в соответствии с универсальным генетическим кодом).

Чем обусловлена столь разная интерпретация рибонуклеотидного контекста для кодонов AUG и GUG? РИС. 8.14 иллюстрирует решающую роль рибосомы в связке со вспомогательными факторами.

Когда малая (30S) субъединица будущей рибосомы связывает мРНК, образуя с ней комплекс инициации, инициирующий кодон ложится в сайт P (который будет сформирован полностью лишь при вступлении большой субъединицы в комплекс инициации). Единственной аминоацил-тРНК, способной вступить в комплекс инициации раньше большой субъединицы, является fMet-тРНК$_f$. Только она может связать неполный сайт P на 30S субъединице, чтобы распознать свой кодон.

В момент, когда к комплексу инициации присоединяется большая субъединица, неполные сайты P и A преобразуются в полностью сформированные сайты. Теперь инициаторная fMet-тРНК$_f$, сохранив свое положение, занимает функционально полноценный сайт P, а сайт A ждет встречи со следующей аминоацил-тРНК, соответствующей второму кодону. В результате взаимодействия двух тРНК образуется первая пептидная связь и начинается элонгация.

Итак, в прочтении конкретного кодона превалирует инициация, если этот кодон AUG (или GUG) находится в пределах сайта связывания рибосомы, потому что ничто, кроме инициаторной fMet-тРНК$_f$, не может встать в неполный сайт P, образуемый при связывании 30S субъединицы с мРНК *de novo*. В рибосоме, продолжающей трансляцию мРНК, имеет место внутреннее прочтение, потому что войти в полный сайт A, образованный обе-

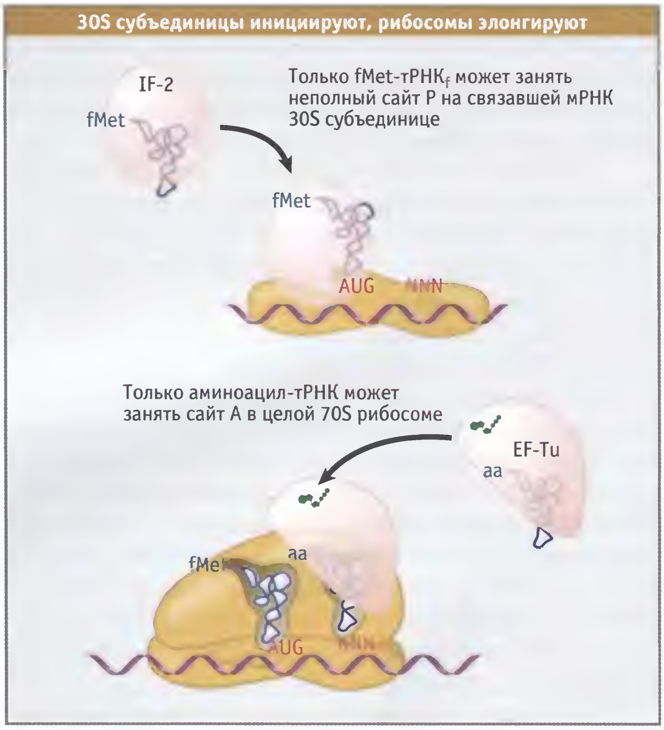


РИС. 8.14 Только fMet-тРНК$_f$ может быть использована 30S субъединицей для инициации. Другие аминоацил-тРНК (аа-тРНК) предназначены для элонгации в 70S рибосоме

ими субъединицами рибосомы, могут только обычные аминоацил-тРНК, но не инициаторная.

В использовании разных видов аминоацил-тРНК очень важна роль вспомогательных факторов. Все аминоацил-тРНК взаимодействуют с рибосомой путем объединения с вспомогательным фактором. При инициации таким фактором служит IF-2 (см. разд. 8.4 «В инициации синтеза белка у бактерий участвуют 30S субъединицы и вспомогательные факторы»), в элонгации — фактор EF-Tu (см. разд. 8.10 «Фактор элонгации EF-Tu загружает аминоацил-тРНК в сайт А»).

Фактор инициации IF-2 помещает инициаторную тРНК в неполный сайт Р. Образуя комплекс с fMet-тРНК$_f$, IF-2 именно в силу своей специфичности отстраняет остальные (регулярные) аминоацил-тРНК от участия в инициации. И наоборот: EF-Tu, который помещает аминоацил-тРНК в сайт А, совсем не умеет связывать fMet-тРНК$_f$ во время элонгации, что исключает ее использование.

Дополнительную идентификацию инициаторной тРНК проводит фактор IF-3, который стабилизирует связывание инициаторной тРНК с мРНК. Фактор IF-3 узнает второе и третье основания кодона AUG только в случае их правильного спаривания с основаниями антикодона.

На **РИС. 8.15** разобрана последовательность событий помещения fMet-тРНК$_f$ в сайт Р фактором IF-2. Сначала IF-2 вкупе с ГТФ связывается с неполным сайтом Р на 30S субъединице рибосомы. В этот момент все факторы инициации оказываются ассоциированными с 30S субъединицей; затем fMet-тРНК$_f$ связывается с IF-2, который тут же помещает ее в неполный сайт Р.

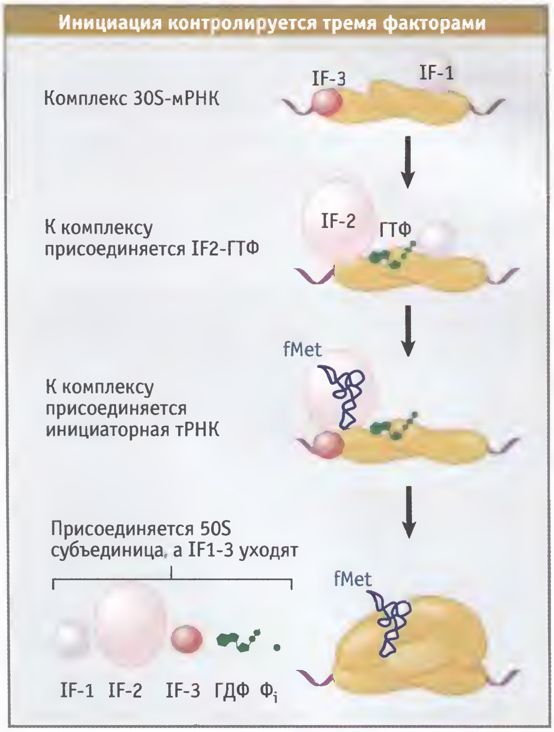


РИС. 8.15 IF-2 нужен для того, чтобы присоединить fMet-тРНК$_f$ к комплексу 30S-мРНК. После того как к комплексу присоединится 50S субъединица, все факторы IF уходят, а ГТФ расщепляется

8.7 Взаимодействие комплементарных участков мРНК и рРНК помогает инициации

Ключевые понятия:

- Последовательность Шайна–Дальгарно, полипуриновый шестичленник в составе бактериальной мРНК, предшествует инициирующему кодону AUG и отделена от него участком в 10 оснований.
- рРНК 30S субъединицы бактериальной рибосомы содержит элемент, комплементарный последовательности Шайна-Дальгарно и взаимодействующий с ней в ходе инициации.

Каждая мРНК содержит множество триплетов AUG. На чем основано опознание одного инициаторного кодона в качестве стартовой точки трансляции? Сайты мРНК, в которых начинается синтез белка, можно идентифицировать экспериментально. Для этого надо запустить трансляцию в условиях, запрещающих элонгацию полипептида. Рибосома, которая не может удлинять полипептид, застревает в сайте инициации. Если к такому заблокированному комплексу инициации добавить рибонуклеазу, она полностью уничтожит всю мРНК, за исключением участка, закрытого рибосомой. Как показано на **РИС. 8.16**, такие фрагменты мРНК можно физически выделить и охарактеризовать.

Фрагменты (фактически это сайты связывания с рибосомой) имеют длину ~30 осн. и обладают двумя свойствами, общими для различных бактериальных мРНК:

- Внутри фрагмента всегда находится инициирующий кодон AUG (реже GUG или UUG).
- Примерно десятью основаниями выше этого AUG находится последовательность, совпадающая (хотя бы частично) с шестичленником

5′... A
G G A
G G
... 3′

Этот полипуриновый шестичленник называют последовательность **Шайна–Дальгарно**. Он комплементарен очень консервативной последовательности

3′...
U C C
U C C
... 5′

содержащейся в 16S рРНК недалеко от 3′-конца. (Степень комплементарности различна для отдельных мРНК и может варьировать от базовой последовательности из 4 оснований (GAGG) до участка в целых 9 оснований, выступающих с обеих сторон за пределы гексамера Шайна–Дальгарно.)

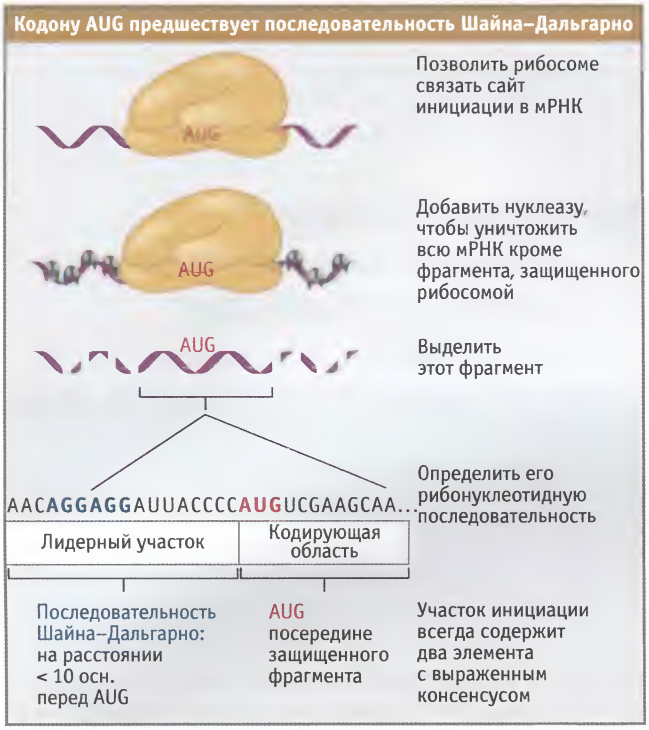


РИС. 8.16 Сайты мРНК, осуществляющие первичный контакт с рибосомой, можно отделить с комплексов инициации *in vitro*. Эти сайты содержат инициирующий кодон AUG, которому предшествует последовательность Шайна–Дальгарно

Взаимодействует ли последовательность Шайна–Дальгарно с комплементарной в рРНК во время связывания мРНК с малой субъединицей рибосомы? По-видимому, да, поскольку мутации в любой из этих последовательностей заметно влияют на ход инициации. Точечные мутации в последовательности Шайна–Дальгарно могут полностью перекрыть трансляцию конкретного типа мРНК. Внесение мутаций в соответствующий участок 16S рРНК будет касаться уже всех мРНК и может представлять опасность для клетки (за счет системных изменений в паттерне синтеза белков). Еще одним убедительным подтверждением взаимодействия является супрессия мутации в последовательности Шайна–Дальгарно (произвольной мРНК) мутацией в 16S рРНК, которая восстанавливает комплементарность соответствующих участков двух молекул.

В малой субъединице рибосомы 3'-концевая область рРНК довольно консервативна: ее последовательности у всех прокариот и эукариот довольно похожи. За одним исключением: у всех эукариот отсутствует сегмент CCUCC (почти целый антипод последовательности Шайна–Дальгарно). По всей видимости, в эукариотической инициации не происходит никакого спаривания между мРНК и 18S рРНК. Значит, эукариотическая инициация по своему механизму довольно сильно отличается от таковой у прокариот.

У бактерий 30S субъединица напрямую связывается с сайтом связывания рибосомы. В результате инициаторный комплекс образуется прямо на участке мРНК, в котором содержится инициирующий кодон AUG. В полицистронных мРНК каждая кодирующая область, как правило, начинается своим собственным сайтом связывания рибосомы.

Гены бактерий экспрессируются таким образом, что трансляция мРНК охватывает цистроны в порядке их следования друг за другом. Не дожидаясь окончания транскрипции, рибосомы начинают осваивать первую кодирующую область новой мРНК. К моменту появления второго сайта связывания рибосомы (во втором цистроне мРНК), в первом цистроне уже полным ходом идет трансляция.

Взаимодействие между кодирующими областями для каждой мРНК свое. Скорее всего, в большинстве случаев рибосомы, независимо друг от друга, складываются из больших и малых субъединиц в начале каждого нового цистрона. Стандартная последовательность событий изображена на **РИС. 8.17**. Когда в конце первого цистрона терминируется синтез первого белка, рибосомы сходят с мРНК и диссоциируют на субъединицы. Новые рибосомы, которые будут транслировать следующую кодирующую область во втором цистроне, должны сложиться из субъединиц в ее собственном, втором, сайте инициации, и так далее.

У некоторых бактериальных мРНК акты трансляции смежных цистронов могут быть сцеплены. Так бывает, когда рибосомы, закончив трансляцию в первом цистроне, автоматически получают доступ к инициирующему кодону во втором цистроне, то есть если промежуток между кодирующими областями невелик, а локальная концентрация рибосом достаточно высока. Интактная рибосома покрывает ~30 оснований мРНК, поэтому она может одновременно контактировать с терминирующим кодоном и со следующим сайтом инициации.



РИС. 8.17 В каждом цистроне полицистронной мРНК имеет место независимое событие инициации, особенно если размер межцистронной области превышает длину участка, покрываемого рибосомой. За диссоциацией рибосом в сайте терминации одного цистрона следует независимая инициация трансляции следующего цистрона

8.8 У эукариот малые субъединицы рибосом сканируют мРНК в поисках сайтов инициации

Ключевые понятия:

- Эукариотическая 40S субъединица рибосомы связывает 5′-конец мРНК и мигрирует вдоль мРНК до обнаружения сайта инициации.
- Эукариотический сайт инициации состоит из десятинуклеотидной последовательности, содержащей кодон AUG.
- 60S субъединицы рибосом присоединяются к комплексу в сайте инициации.

Инициация синтеза белка в цитоплазме эукариотической клетки в целом похожа на инициацию у бактерий, но порядок событий немного отличается, а число вспомогательных факторов больше. При более подробном рассмотрении следует особо отметить различия в способе стыковки малых субъединиц рибосомы со своими сайтами связывания в мРНК. Малые 40S субъединицы эукариотических рибосом сначала находят 5′-конец мРНК, затем перемещаются оттуда в сайт инициации; там они соединяются с большими субъединицами. У прокариот малые субъединицы связываются непосредственно с сайтом инициации, который, таким образом, является и сайтом первичного контакта мРНК с малой субъединицей, и сайтом сборки рибосомы (см. разд. 8.4 «В инициации синтеза белка у бактерий участвуют 30S субъединицы и вспомогательные факторы»).

Практически все эукариотические мРНК являются моноцистронными. При этом они, как правило, длинные — значительно длиннее своей кодирующей области. Большинство эукариотических мРНК имеет длину 1000–2000 осн., метилированный кэп на 5′-конце, и 100–200 оснований поли(А) на 3′-конце.

Длина 5′-концевой нетранслируемой области, так называемого лидера, обычно не превышает 100 осн. Длина кодирующей области соответствует размеру белка. Длина 3′-концевой нетранслируемой области, так называемого трейлера, в среднем составляет 100 осн.

Трансляция эукариотической мРНК начинается с ее опознавания. Главной детерминантой в этом опознавании служит метилированный кэп, которым маркирован 5′-конец; мРНК, у которой удален кэп, не будет эффективно транслироваться *in vitro*. Для связывания 40S субъединицы с мРНК требуется несколько факторов инициации; некоторые из этих факторов распознают кэп по его структурным особенностям.

Модификация 5′-конца, так называемое кэпирование, происходит почти со всеми клеточными (а также вирусными) мРНК. Наличие кэпов на 5′-концах мРНК очень существенно для их трансляции в цитоплазме (но не в органеллах; там кэпы не нужны). Из этого правила есть одно исключение: мРНК некоторых вирусов (таких как полиовирус) не имеют кэпов и, тем не менее, хорошо транслируются *in vitro*. Для них работает альтернативный путь инициации трансляции, где кэп не нужен.

Некоторые вирусы используют это различие в качестве адаптационного преимущества. В развитии полиовирусной инфекции происходит ингибирование трансляции кэпированных мРНК клетки-хозяина. Это ингибирование осуществляется через воздействие на кэп-связывающие белки — факторы инициации трансляции клеточных мРНК, которые не нужны для трансляции некэпированных вирусных мРНК.

Мы привыкли смотреть на процесс инициации так, будто бы сайт связывания рибосомы всегда находится в свободном доступе. На самом деле этот сайт запросто может быть «спрятан» во вторичной структуре мРНК. Посадка рибосомы на мРНК требует присутствия нескольких дополнительных факторов; важной частью их функции является устранение вторичной структуры мРНК (см. рис. 8.20).

Если инициирующий кодон AUG находится в пределах 40 осн. от 5′-конца мРНК, одна 40S субъединица может одновременно покрывать и кэп, и AUG. Во многих мРНК кэп с AUG разнесены сильнее; в отдельных случаях расстояние между кэпом и AUG достигает 1000 осн. Но и в этих экстремальных случаях для образования стабильного комплекса инициации необходимо наличие кэпа на 5′-конце мРНК. Как может рибосома во время инициации опираться на точки, столь отдаленные друг от друга?

Для ответа на этот вопрос существует так называемая модель сканирования, которая разобрана на РИС. 8.18. Предполагается, что 40S субъединица сначала узнает кэп,

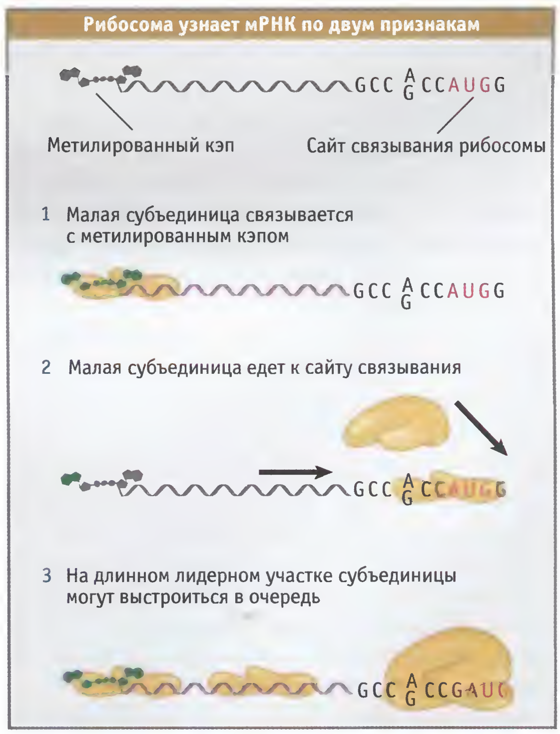


РИС. 8.18 Малые субъединицы эукариотических рибосом мигрируют с 5′-конца мРНК к сайту связывания рибосомы, в котором находится инициаторный кодон AUG

а затем начинает двигаться вдоль мРНК, как бы изучая, сканируя ее на предмет сайта инициации. Сканирование, от 5′-конца вглубь мРНК, представляет собой линейный процесс. Когда 40S субъединица сканирует лидерный участок, она может справиться со шпилькой вторичной структуры, если ее стабильность не превышает 30 ккал. Более стабильные шпильки мешают миграции, поскольку 40S субъединице трудно их расплетать.

Миграция останавливается в момент, когда 40S субъединица встречает инициаторный кодон AUG. Чаще всего, первый же триплет AUG на пути субъединицы как раз и оказывается инициаторным кодоном. Самого по себе триплета AUG недостаточно для того, чтобы 40S субъединица остановилась для инициации; триплет AUG должен находиться в правильном контексте. Важнейшими детерминантами контекста являются азотистые основания в положениях –4 и +1. Инициаторный кодон подлежит узнаванию только в контексте NNNPuNN*AUG*G. Пурин (Pu; A или G), стоящий тремя положениями выше AUG, а также G сразу за AUG на порядок повышают эффективность трансляции. Если мРНК имеет протяженный лидерный участок, ее 5′-конец будет освобождаться для опознания следующими малыми субъединицами еще до того, как первая рибосома, начав синтез, покинет сайт инициации. В этом случае лидерный участок заполняет целая очередь малых рибосомных субъединиц, мигрирующих от 5′-конца к сайту инициации.

В 5′-концевой нетранслируемой области многих мРНК можно обнаружить рибонуклеотиды A, U и G, идущие подряд. (В мРНК, которые транслируются наиболее эффективно, такого нет; инициирующий кодон всегда соответствует самому первому триплету AUG.) Возникает вопрос: когда мигрирующая малая субъединица встречает «неинициаторный» AUG в 5′-концевой нетранслируемой области, почему она не принимает его за старт-кодон? Во-первых, этот AUG не имеет нужного контекста. Во-вторых, в редких случаях трансляция с него все-таки идет, только эта бессмысленная трансляция очень быстро терминируется, и рибосома возобновляет сканирование мРНК для встречи с настоящим инициаторным AUG.

Абсолютному большинству эукариотических инициаций предшествует сканирование мРНК, начиная от кэпа на 5′-конце. Есть еще альтернативный способ инициации, при котором 40S субъединица сразу занимает внутренний сайт мРНК, называемый IRES. При таком способе инициации, характерном для мРНК некоторых вирусов, полностью решается проблема наличия триплетов AUG в 5′-концевой нетранслируемой области мРНК. Элементов IRES идентифицировано несколько, и они мало похожи между собой. По особенностям взаимодействия с 40S субъединицей можно выделить три типа IRES:

- IRES первого типа содержит инициаторный кодон AUG возле своей верхней границы (то есть ближней к 5′-концу мРНК). 40S субъединица связывается прямо с инициаторным кодоном. Связыванию помогает тот же набор факторов, что и при стандартной инициации (когда 40S субъединица едет к инициаторному AUG с самого 5′-конца).
- IRES второго типа находится на 100 осн. выше (то есть ближе к 5′-концу), чем инициаторный AUG; после связывания с таким IRES субъединица все-таки должна мигрировать, по всей видимости, сканируя мРНК — так же, как при стандартной инициации.
- IRES вируса гепатита C, единственный в своем роде, умеет связывать 40S субъединицу без поддержки факторов инициации. У такой инициации свой, совершенно особый порядок событий. Комплекс 40S-мРНК, образующийся в этом случае, также содержит факторы инициации и инициаторную тРНК.

Использование IRES особенно актуально для пикорнавирусов (IRES были впервые описаны именно у них). Во время пикорнавирусной инфекции происходит ингибирование синтеза всех клеточных белков: вирус уничтожает кэпы и ингибирует кэп-связывающие факторы инициации (см. разд. 8.9 «У эукариот множество факторов инициации объединены в комплекс»).

Независимо от способа, по которому идет инициация, в сайте инициации ассоциация малой субъединицы с мРНК окончательно стабилизируется. Малая субъединица в одиночку закрывает от рибонуклеазы довольно длинный участок мРНК, ~60 осн. После соединения малой и большой субъединиц участок, закрываемый теперь уже целой эукариотической рибосомой, сокращается до 30–40 осн. (Целая прокариотическая рибосома закрывает собой участок примерно такой же длины.)

8.9 У эукариот множество факторов инициации объединено в комплекс

Ключевые понятия:

- Факторы инициации необходимы на всех стадиях инициации, включая связывание инициаторной тРНК, первичного связывания 40S субъединицы рибосомы с мРНК, ее движения вдоль мРНК, а также для присоединения 60S субъединицы.
- Эукариотическая инициаторная тРНК — это Met-тРНК, отличающаяся от Met-тРНК, используемой при элонгации, ее метионин не формилирован.
- Фактор инициации eIF2, ассоциированный с ГТФ, связывает инициаторную Met-тРНК$_i$; весь комплекс связывается с 40S субъединицей еще до того, как она свяжется с мРНК.

Инициация синтеза белка в цитоплазме эукариотической клетки в общих чертах обнаруживает явное сходство с инициацией бактерий. Тем не менее, у двух инициаций есть явное сходство; оно проявляется, например, в использовании специального инициаторного кодона и особой инициаторной тРНК. В эукариотической инициации всегда используется инициаторный

кодон AUG. он соответствует аминокислоте метионину. Инициаторная тРНК (тРНК$_i^{Met}$) отличается от регулярных аминоацил-тРНК, а ее метионин не подлежит формилированию. Так как отличия между инициаторной и элонгаторной Met-тРНК приходятся исключительно на долю тРНК, то для их различения применяют индексы: Met-тРНК$_i$ — для инициаторной и Met-тРНК$_m$ — для используемой при элонгации.

У тРНК$_i^{Met}$ дрожжей есть как минимум две уникальных особенности: нестандартная пространственная структура и наличие фосфатной группы на 2′-рибозе шестьдесят четвертого рибонуклеотида. Если запретить это фосфорилирование, то «заряженная» метионином инициаторная тРНК сможет участвовать и в элонгации. Таким образом, эукариоты используют иной принцип разделения Met-тРНК на функциональные пулы (для сравнения см. рис. 8.13).

Эукариотические клетки располагают бóльшим числом факторов инициации, чем бактериальные. На сегодняшний день изучено 12 факторов, которые, прямо или косвенно, определяют ход эукариотической инициации. Им присвоили те же названия, что и бактериальным факторам (иногда с соблюдением функциональных и структурных аналогий там, где это было возможно; с добавлением префикса «e», обозначающего их принадлежность эукариотам: eukaryotic). Участие факторов критично во всех аспектах инициации, включая:

- формирование комплекса инициации с 5′-концом мРНК;
- формирование комплекса с Met-тРНК$_i$;
- связывание комплексов мРНК-факторы и Met-тРНК$_i$-факторы;
- возможность сканирования мРНК рибосомой с 5′-конца до первого AUG;
- обнаружение связи инициаторной тРНК с AUG в сайтах инициации;
- способствование прикреплению 60S субъединицы.

На **РИС. 8.19** выстроена последовательность всех промежуточных этапов инициации, с указанием факторов, участвующих на каждом из них. Факторы eIF2 и eIF3 связываются с 40S субъединицей рибосомы; факторы eIF4A, eIF4B и eIF4F связываются с мРНК; факторы eIF1 и eIF1A связываются с комплексом малой субъединицы и мРНК.

На **РИС. 8.20** изображена группа факторов, связывающих 5′-конец мРНК. Фактор eIF4F, как видно из рисунка, представляет собой белковый комплекс, состоящий из трех факторов инициации. Не вполне ясно, существует он как комплекс до присоединения к мРНК или отдельные субъединицы собираются вместе прямо на мРНК. В состав eIF4F входят: кэп-связывающая субъединица eIF4E, хеликаза eIF4A и «распорка» eIF4G (scaffolding protein). После того как eIF4E связывает кэп, eIF4A расплетает всю вторичную структуру мРНК на участке из первых 15 оснований. Энергию для этой хеликазной активности дает гидролиз АТФ. Дальнейшее расплетание вторичной структуры выполняется все тем же eIF4A, но уже вместе с другим фактором, eIF4B. eIF4G поддерживает целостность всего комплекса инициации, связывая между собой его компоненты.

Субъединица eIF4E служит мишенью для регуляции. Ее активность может быть увеличена фосфорилированием. Различные сигнальные пути, подавляющие либо

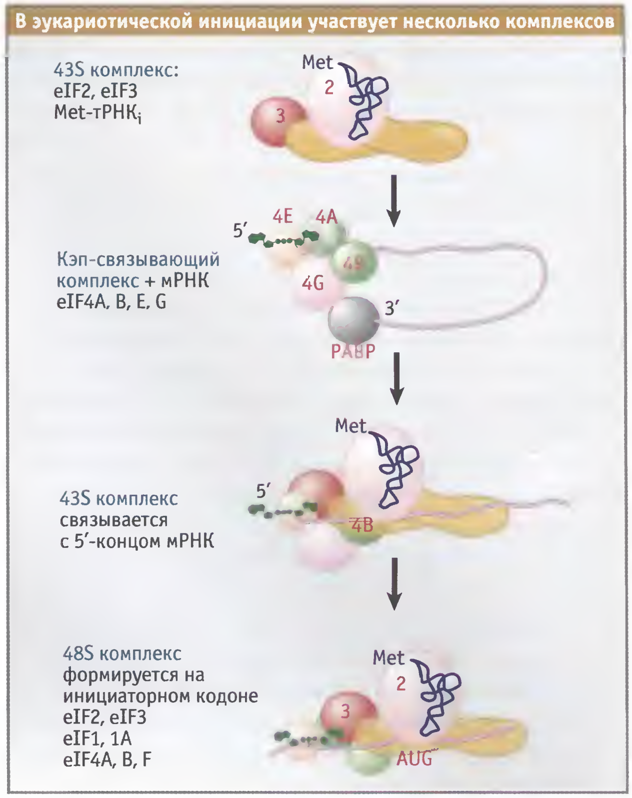


РИС. 8.19 Некоторые факторы инициации связываются с 40S субъединицей рибосомы и вместе с ней образуют 43S комплекс. Другие факторы инициации связываются с мРНК. 43S комплекс движется вдоль мРНК; на инициаторном кодоне он преобразуется в 48S комплекс, который также поддается выделению

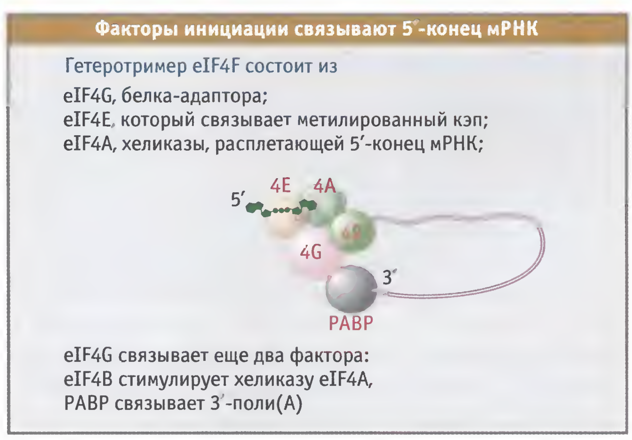


РИС. 8.20 Гетеротример eIF4F связывает 5′-конец мРНК, а также последующие факторы

интенсифицирующие синтез белка, срабатывают на запуск, соответственно, фосфорилирования либо дефосфорилирования eIF4E. Кроме того, доступность eIF4E контролируется через его связывание белками-ингибиторами. Эти белки (4E-BP1, -2 и -3), связывая eIF4E, не дают ему участвовать в инициации. Субъединица eIF4G также может служить мишенью для регуляции: она подлежит деградации при пикорнавирусной инфекции (когда вирус пытается мешать инициации трансляции клеточных мРНК с кэпа на 5′-конце, см. разд. 8.8 «В эукариотической инициации малые субъединицы рибосом сканируют мРНК в поисках инициирующего кодона»).

Наличие поли(А) на 3′-конце мРНК способствует образованию комплекса инициации на ее 5′-конце. Специальный белок (Pab1p у дрожжей) одновременно связывает поли(А) и eIF4G («распорку» комплекса инициации). При этом мРНК неизбежно обретает вид кольца — ведь оба ее конца, 5′ и 3′, теперь удерживаются одним белковым комплексом (см. рис. 8.20). В чем смысл образования такой замкнутой петли, точно неизвестно; перечислим лишь возможные эффекты:

- Стимуляция инициации или трансляции.
- Рибосомы как бы циркулируют по мРНК: после терминации возле 3′-конца освободившиеся рибосомные субъединицы уже находятся близко к 5′-концу и снова готовы к инициации.
- мРНК обретает дополнительную устойчивость к деградации.
- Факторы, ассоциированные с 3′-концом, имеют возможность регулировать инициацию трансляции.

Субъединица eIF2 — ключевой фактор в присоединении Met-тРНК$_i$. Фактор eIF2 — это типичный мономерный G-белок (ГТФаза), который активен, когда укомплектован ГТФ, и неактивен, когда связан с ГДФ (после гидролиза ГТФ с выходом энергии). На РИС. 8.21 показано, что eIF2-ГТФ связывается с Met-тРНК$_i$. Продукт связывания иногда называют тройственным комплексом (ternary complex), потому что он состоит из трех компонентов (eIF2, ГТФ и Met-тРНК$_i$).

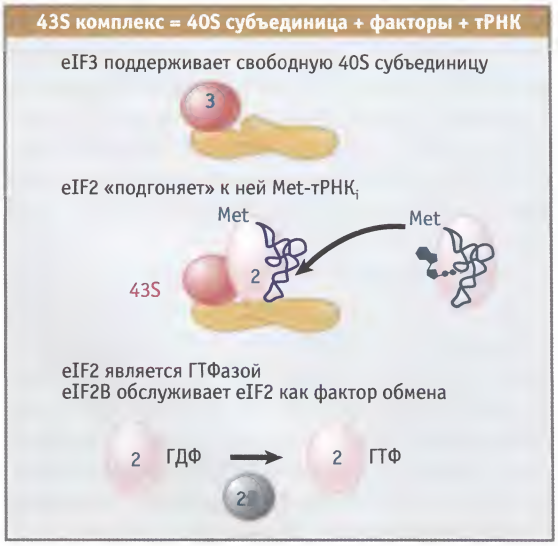


РИС. 8.22 Факторы инициации соединяют инициаторную Met-тРНК с 40S субъединицей рибосомы, образуя 43S комплекс. Далее по ходу реакции ГТФ гидролизуется, и eIF2 высвобождается в форме eIF2-ГДФ. Фактор обмена eIF2B восстанавливает активную форму eIF2



РИС. 8.21 При инициации трансляции у эукариот фактор eIF2 и ГТФ образуют тройственный комплекс с Met-тРНК$_i$. Тройственный комплекс связывает свободную 40S субъединицу, которая прикрепляется к 5′-концу мРНК. Затем ГТФ гидролизуется, и eIF2 высвобождается в форме eIF2-ГДФ. Активную форму eIF2 восстанавливает фактор eIF2B

На РИС. 8.22 показано, что тройственный комплекс приводит Met-тРНК$_i$ в контакт с малой субъединицей рибосомы. Фактически он загружает Met-тРНК$_i$ на 40S субъединицу: получается комплекс инициации 43S. Для этой реакции не нужна мРНК. Напротив, для того чтобы малая субъединица рибосомы смогла связать мРНК, необходимо присутствие инициаторной тРНК Met-тРНК$_i$. В 43S комплексе присутствует еще один фактор, eIF3, который поддерживает свободное состояние малой субъединицы (препятствует преждевременной ассоциации большой и малой субъединиц). На самом деле, eIF3 — это целый комплекс, насчитывающий от восьми до десяти белковых субъединиц.

Следующим шагом 43S комплекс связывает 5′-конец мРНК. На РИС. 8.23 показано, что белок-белковые взаимодействия, обусловливающие этот этап инициации, еще не до конца изучены. Вероятно, наряду с малой субъединицей рибосомы и мРНК, в них участвуют eIF4G и eIF3, причем eIF4G связывает eIF3, и это взаимодействие служит мостом между eIF4G и малой субъединицей: теперь они образуют единый массив. По сути, eIF4G рекрутирует малую субъединицу — подтягивает ее ко всему eIF4F, а через него — к мРНК.

После того как малая субъединица связала мРНК, она движется по ней к ближайшему (чаще всего) кодо-



РИС. 8.23 Вступление мРНК в 43S комплекс обусловлено взаимодействиями, в которых участвуют факторы инициации

ну AUG. Движение требует затрат энергии в форме АТФ и происходит при поддержке факторов eIF1 и eIF1A. На РИС. 8.24 показано, что малая субъединица останавливается, как только достигнет сайта инициации, превращаясь в 48S комплекс.

Присоединение 60S субъединицы рибосомы к инициаторному комплексу не может состояться до тех пор, пока eIF2 и eIF3 не покинут комплекс. Им помогает уйти еще один фактор, eIF5; высвобождаясь из комплекса, eIF2 гидролизует свой ГТФ (до ГДФ). Реакция имеет место на малой субъединице; ее необходимым условием является взаимодействие антикодона инициаторной тРНК с комплементарным инициаторным кодоном AUG. Все остальные факторы инициации высвобождаются уже после образования полной 80S рибосомы.

Когда, в конце концов, фактор eIF5B помогает 60S субъединице присоединиться к инициаторному комплексу, полная 80S рибосома готова к элонгации. Фактор eIF5B похож на прокариотический фактор IF2 (который также проявляет ГТФазную активность — в дополнение к важной роли связывания инициаторной тРНК).

После того как факторы инициации покидают рибосому, они подлежат повторному использованию в последующих циклах инициации. Фактор eIF2, гидролизовавший свой ГТФ, нуждается в регенерации (восстановлении активной формы). Регенерацию eIF2 выполняет еще один фактор инициации, eIF2B. Он вытесняет ГДФ из комплекса с eIF2, на место ГДФ приходит ГТФ. Формально, eIF2B можно охарактеризовать как фактор обмена (exchange factor) при ГТФазе eIF2.

Фактор eIF2 служит мишенью для регуляции. Он является потенциальным субстратом нескольких регуляторных киназ — эффекторов сигнальных путей. Они фосфорилируют α-субъединицу eIF2, что не позволяет фактору обмена eIF2B восстанавливать активную форму eIF2. Таким образом, фосфорилирование eIF2 по α-субъединице быстро приводит к нехватке активного eIF2B и общему ингибированию синтеза белка.

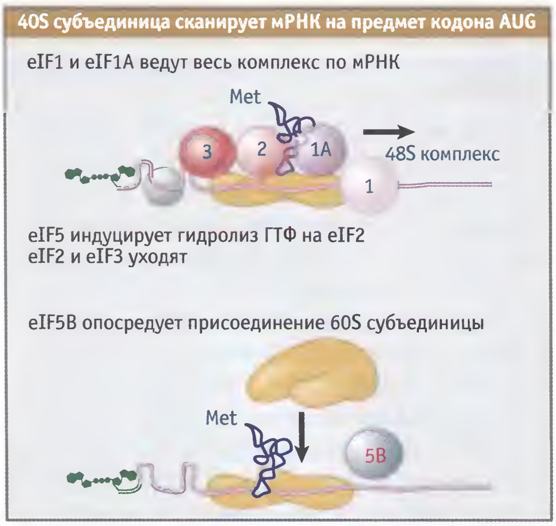


РИС. 8.24 eIF1 и eIF1A помогают инициаторному комплексу 43S сканировать мРНК до тех пор, пока он не достигнет кодона AUG. Гидролиз ГТФ на eIF2 приводит к уходу eIF2 вместе с eIF3. Фактор eIF5B опосредует воссоединение большой и малой субъединиц рибосомы

8.10 Фактор элонгации EF-Tu загружает аминоацил-тРНК в сайт А

Ключевые понятия:

- EF-Tu представляет собой мономерный G-белок; его активная форма (укомплектованная ГТФ) связывает аминоацил-тРНК.
- Комплекс EF-Tu–ГТФ–аминоацил-тРНК взаимодействует с сайтом А рибосомы.

Как только на инициирующем кодоне бактериальной мРНК сформировалась полная рибосома, система готова совершить первый рабочий цикл элонгации: аминоацил-тРНК войдет в сайт А и провзаимодействует с пептидил-тРНК, занимающей сайт Р. (В первом рабочем цикле элонгации вместо пептидил-тРНК будет участвовать fMet-тРНК$_f$.) Войти в сайт А может любая аминоацил-тРНК, за исключением инициаторной. Проводником ей служит специальный белок — **фактор элонгации**. Бактериальный фактор элонгации называют **EF-Tu** (сразу отметим, что эукариотический фактор элонгации работает аналогично бактериальному). EF-Tu имеет очень консервативную структуру: факторы элонгации всех бактерий и митохондрий не только очень похожи между собой, но еще и гомологичны эукариотическим факторам элонгации.

Подобно своему аналогу, участвующему в инициации (фактору IF-2), фактор EF-Tu ассоциирован с рибосомой только непосредственно «при исполнении», то есть когда он приносит на рибосому новую аминоацил-

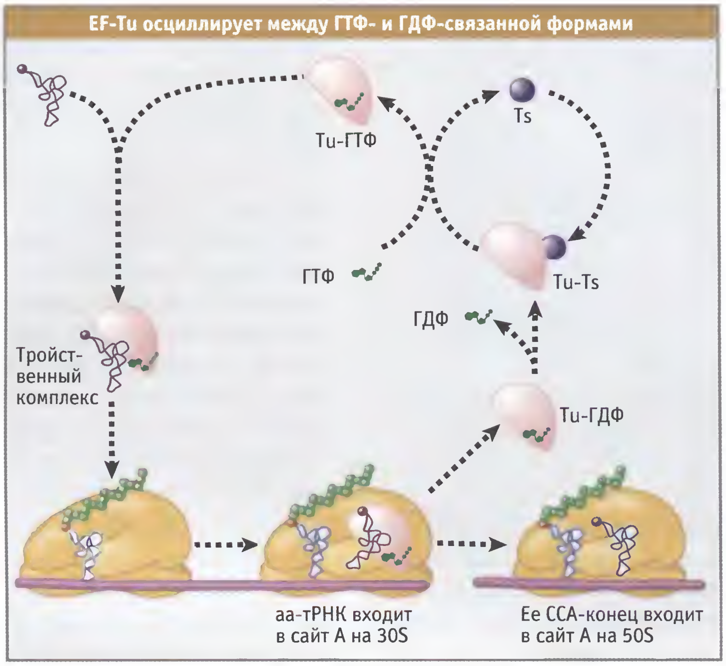


РИС. 8.25 EF-Tu–ГТФ помещает аминоацил-тРНК в рибосому и покидает рибосому уже как EF-Tu–ГДФ. Чтобы заменить этот ГДФ на ГТФ, требуется EF-Ts. EF-Tu–ГТФ может рекрутировать любую аминоацил-тРНК, за исключением fMet-тРНК$_f$. Поэтому fMet-тРНК$_f$ не участвует в трансляции внутренних кодонов AUG и GUG

тРНК. Как только аминоацил-тРНК занимает свое место (сайт А), EF-Tu покидает рибосому, чтобы начать работать с другой аминоацил-тРНК. Таким образом, EF-Tu циклически ассоциирует и диссоциирует с рибосомой (на самом деле таков режим работы всех вспомогательных факторов).

Схема циклических поставок аминоацил-тРНК в рибосому изображена на **РИС. 8.25**. Фактор EF-Tu представляет собой мономерный G-белок. Он несет на себе гуанин-нуклеотид (ГТФ либо ГДФ). Активность фактора определяется состоянием гуанин-нуклеотида и контролируется через него:

- Если в наличии ГТФ, фактор находится в активном состоянии.
- При гидролизе ГТФ до ГДФ фактор инактивируется.
- При замене ГДФ на ГТФ активное состояние восстанавливается.

Бинарный комплекс EF-Tu с ГТФ встречает аминоацил-тРНК и связывает ее: образуется тройственный комплекс EF-Tu–ГТФ-аминоацил-тРНК. Тройственный комплекс специфичен к сайтам А только тех рибосом, у которых в сайте Р находится пептидил-тРНК. Поэтому само связывание тройственного комплекса с рибосомой можно считать гарантией того, что аминоацил-тРНК и пептидил-тРНК заняли правильные исходные положения для образования новой пептидной связи.

Загрузка аминоацил-тРНК в сайт А происходит в две ступени. Сначала антикодон тРНК заходит в сайт А (точнее, в ту часть сайта А, которая принадлежит 30S субъединице). Там антикодон встречается с кодоном, и если они узнают друг друга, это приводит к изменению конформации рибосомы. Это стабилизирует связывание тРНК и заставляет EF-Tu гидролизовать свой ГТФ. Затем другой конец тРНК (CCA-шпилька, противоположная антикодону и несущая аминоацил) встает в свою часть сайта А, принадлежащую 50S субъединице. Бинарный комплекс EF-Tu–ГДФ покидает рибосому; в такой форме EF-Tu неактивен (то есть не может эффективно связывать аминоацил-тРНК).

Для регенерации (восстановления активной формы) EF-Tu в целях повторного использования нужен еще один фактор элонгации, EF-Ts. Он служит G-белку EF-Tu фактором обмена. Сначала EF-Ts вытесняет ГДФ из комплекса EF-Tu–ГДФ; получается комбинированный фактор EF-Tu–EF-Ts. Затем ГТФ вытесняет EF-Ts и встает на его место: получается активный бинарный комплекс EF-Tu–ГТФ. Он связывает аминоацил-тРНК и несет ее на рибосому; освободившийся EF-Ts уходит помогать регенерации других использованных EF-Tu.

В каждой бактерии есть ~70 000 молекул EF-Tu (они составляют ~5% общей массы белка). Число молекул EF-Tu сопоставимо с числом молекул аминоацил-тРНК. Скорее всего, большинство клеточных аминоацил-тРНК находится в составе тройственных комплексов. Молекулы EF-Ts не столь многочисленны (их насчитывается ~10 000 на клетку; примерно столько же, сколько рибосом). Кинетика взаимодействия EF-Tu с EF-Ts говорит о том, что оно очень не-

устойчиво: комплекс EF-Tu–EF-Ts существует очень недолго. EF-Tu быстро переходит в форму, связанную с ГТФ, а затем — в тройственный комплекс с аминоацил-тРНК.

Роль ГТФ в тройственном комплексе стала предметом отдельного исследования. В серии экспериментов ГТФ заменяли искусственным аналогом, который не может быть гидролизован. Это химическое соединение, **GMP-PCP**, почти полностью идентично ГТФ, только вместо атома кислорода между вторым и третьим фосфатами, P_i^{β} и P_i^{γ} стоит негидролизуемый метиленовый мостик. Оказалось, что тройственный комплекс в этом случае тоже образуется, тоже приводит аминоацил-тРНК в рибосому, где она тоже связывается с сайтом A. Но образования пептидной связи после этого не происходит. По всей видимости, интактная ГТФ нужна для того, чтобы аминоацил-тРНК встала в сайт A; для дальнейшего развития событий необходим гидролиз ГТФ.

Антибиотик **кирромицин** является прямым ингибитором функции EF-Tu. Находясь в комплексе с кирромицином, EF-Tu все еще способен привести аминоацил-тРНК в рибосому и поставить ее в сайт A. Но «отработавший» комплекс EF-Tu–ГДФ уже не может покинуть рибосому — не пускает кирромицин. Длительное присутствие EF-Tu–ГДФ в рибосоме препятствует образованию пептидной связи между пептидил-тРНК и аминоацил-тРНК. В результате синтез белка застопоривается; рибосома прекращает двигаться по мРНК.

Эффект действия кирромицина демонстрирует, как ингибирование одного шага белкового синтеза блокирует последующие. Длительное присутствие EF-Tu–ГДФ на рибосоме мешает аминоацильной части тРНК занять сайт A на 50S субъединице полной рибосомы (см. рис. 8.31). Таким образом, уход EF-Tu–ГДФ необходим рибосоме, чтобы образовать следующую пептидную связь. На самом деле, принцип строгой преемственности событий молекулярного уровня актуален на всех этапах синтеза белка: каждая реакция должна быть полностью завершена, только тогда сможет начаться следующая.

Взаимодействие EF-Tu с рибосомой играет важную роль в контроле «качества воспроизведения» (то есть точности соответствия цепочки аминокислотных остатков растущего полипептида своей матрице — цепочке кодонов мРНК). EF-Tu–ГТФ приводит в рибосому произвольную аминоацил-тРНК: ее соответствие кодону, ждущему в сайте A, заранее никак не предусмотрено (скорее всего, соответствия между ними нет). Гидролиз EF-Tu–ГТФ в рибосоме идет сравнительно медленно, как бы предоставляя время для проверки соответствия. Если аминоацил-тРНК не соответствует кодону в сайте A, она успеет выскочить из него еще до того, как EF-Tu гидролизует свой ГТФ. Если не успеет, ничего страшного: EF-Tu–ГДФ покидает рибосому опять же довольно медленно, так что неподходящая тРНК может покинуть рибосому и после гидролиза. Общий принцип реакций с участием EF-Tu заключается в том, что их темп достаточно замедлен для того, чтобы неподходящие аминоацил-тРНК успевали покидать рибосому прежде, чем их аминоацил будет ошибочно и необратимо включен в полипептид.

В цитоплазме эукариотической клетки роль EF-Tu играет гомологичный фактор eEF1α (тоже высокопредставленный белок). Подобно EF-Tu, он приносит аминоацил-тРНК на рибосому; его функция точно так же требует гидролиза высокоэнергетической связи ГТФ, что сопровождается переходом фактора в неактивную форму. Активность eEF1α восстанавливается фактором eEF1βγ (эукариотическим аналогом EF-Ts).

8.11 Растущий полипептид переносится на аминоацил-тРНК

Ключевые понятия:

- 50S субъединица рибосомы обладает активностью пептидилтрансферазы.
- Насцентный полипептид переносится с пептидил-тРНК в сайте P на аминоацил-тРНК в сайте A.
- Образование пептидной связи приводит к деацилированию тРНК в сайте P и формированию пептидил-тРНК в сайте A.

Пока растущий пептид удлиняется («элонгируется») за счет переноса всей полипептидной цепочки на новую аминоацил-тРНК, из сайта P в сайт A, рибосома стоит на месте. Реакция переноса изображена на РИС. 8.26. Катализ этой реакции можно рассматривать как активность **пептидилтрансферазы**.

Чтобы лучше понять механизм реакции переноса, попробуем проанализировать способность антибиотика **пуромицина** к ингибированию синтеза белка. Молекула пуромицина напоминает аминокислоту, прикрепленную к концевому аденозину тРНК; из РИС. 8.27 видно, что у пуромицина вместо атома кислорода, соединяющего аминокислоту с тРНК, стоит атом азота. Рибосома принимает молекулу пуромицина за аминоацил-тРНК и переносит растущий полипептид на NH_2-группу пуромицина.

Комплекс полипептидил-пуромицин, никак не закрепленный в сайте A, довольно быстро покидает рибосому, унося на себе весь полипептид. Так, по вине пуромицина, происходит преждевременная терминация синтеза белка. Под действием пуромицина клетка, разумеется, гибнет.

Активность пептидилтрансферазы — это собственная активность большой (50S или 60S) субъединицы рибосомы; здесь она работает как фермент. Реакция запускается в тот момент, когда EF-Tu отпускает аминоацил-тРНК в сайте A, и она, как бы покачнувшись, переводит свой аминоацил в положение, близкое к

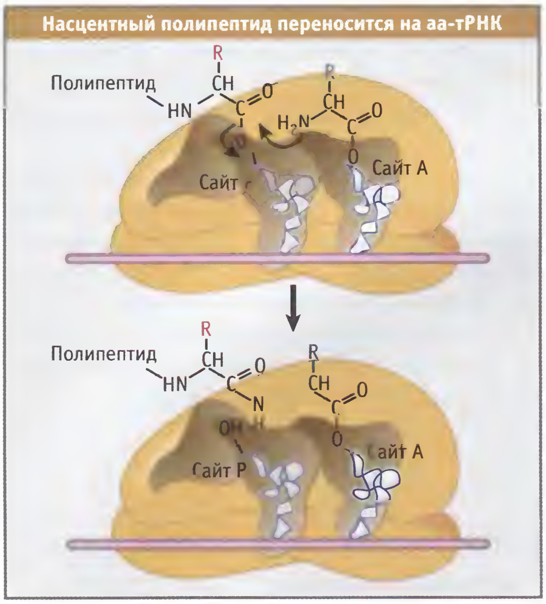


РИС. 8.26 Образование пептидной связи — это реакция между полипептидом пептидил-тРНК, занимающей сайт P, и аминокислотой аминоацил-тРНК, занимающей сайт A

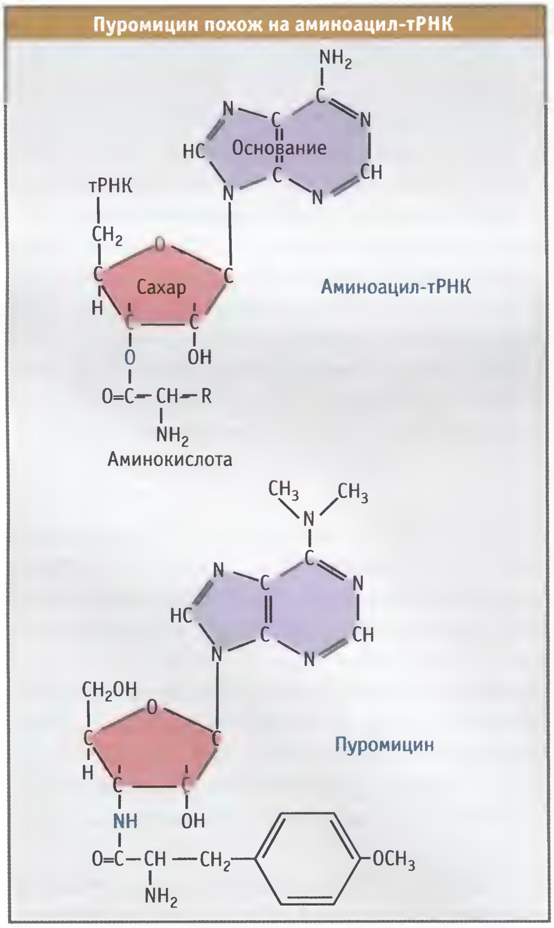


РИС. 8.27 Пуромицин имитирует аминоацил-тРНК, поскольку напоминает ароматическую аминокислоту, пришитую к тандему сахара с азотистым основанием

растущему полипептиду на тРНК в сайте P. На стыке сайтов P и A в большой субъединице рибосомы как раз и находится активный центр пептидилтрансферазы. Реакция происходит быстро, причем акт катализа (по крайней мере, у бактерий) можно целиком отнести на счет большой рРНК 50S субъединицы рибосомы (см. разд. 8.19 «23S рРНК обладает пептидил-трансферазной активностью»).

8.12 При транслокации рибосома приходит в движение

Ключевые понятия:

- При транслокации мРНК протягивается сквозь рибосому на три основания.
- При транслокации деацилированная тРНК переходит в сайт E, а пептидил-тРНК — в сайт P; сразу после транслокации сайт A оказывается не занятым.
- Модель гибридного состояния строится на предположении, что транслокация происходит в два этапа: сначала 50S субъединица рибосомы сдвигается относительно 30S субъединицы, затем 30S субъединица сдвигается относительно мРНК, возвращая рибосоме ее первоначальную конформацию.

Рабочий цикл рибосомы (единичный шаг элонгации, добавление одной аминокислоты к растущему полипептиду) завершается актом транслокации: рибосома продвигается вдоль мРНК на три рибонуклеотида вперед (по направлению к 3'-концу мРНК). На РИС. 8.28 показано, что во время транслокации деацилированная тРНК выдвигается из сайта P «на выход», а новоиспеченная пептидил-тРНК переезжает на своем кодоне мРНК из сайта A в сайт P. После транслокации сайт A оказывается пустым (если не считать мРНК), ожидая следующую аминоацил-тРНК. У бактерий «разряженная» тРНК переходит, как показано на рисунке, из сайта P в сайт E, и оттуда уже окончательно уходит в цитоплазму. На эукариотической рибосоме нет сайта E; отработавшая тРНК уходит из сайта P прямо в цитоплазму. Сайты A и P образованы обеими субъединицами рибосомы; сайт E (у бактерий) образован большой субъединицей рибосомы (хотя малая субъединица с ним тоже контактирует).

Большинство исследователей транслокации склоняются к так называемой «модели гибридного состояния» (hybrid state model). Согласно этой модели, транслокация происходит в два этапа. На РИС. 8.29 показано, что сначала происходит сдвиг 50S субъединицы относительно 30S субъединицы. Второй сдвиг (30S субъединицы относительно мРНК) довершает транслокацию и восстанавливает первоначальную конформацию рибосомы. В основу этой модели легло наблюдение, что паттерн контактов тРНК с рибосомой (снятый при помощи техники химических «следов» — футпринтинга, *англ.* footprinting) изменяется не непрерывно, а в два

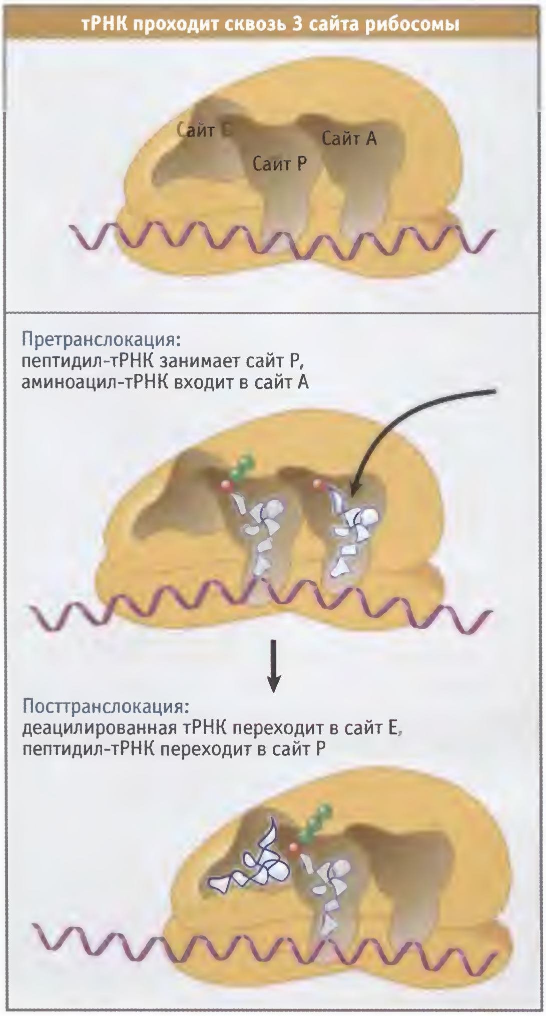


РИС. 8.28 На бактериальной рибосоме есть три сайта связывания тРНК. Аминоацил-тРНК входит в сайт А; при этом сайт Р уже занят пептидил-тРНК. Образование пептидной связи сопряжено с деацилированием тРНК в сайте Р и превращением аминоацил- в пептидил-тРНК в сайте А. При транслокации деацилированная тРНК переходит в сайт Е, а пептидил-тРНК — в сайт Р

этапа. Когда к рибосоме, у которой в сайте Р находится аминоацил-тРНК, присоединяется пуромицин, контакты тРНК переходят в 50S субъединице из сайта Р в сайт Е; при этом контакты тРНК с 30S субъединицей никак не меняются. Значит, после реакции переноса 50S субъединица пришла в движение, а 30S субъединица осталась стоять на месте.

В приложении к процессу синтеза эти данные можно интерпретировать так: сначала аминоацильные концы тРНК (расположенные в 50S субъединице) «переезжают» в новые сайты. В это время антикодоны обеих тРНК еще сохраняют связь со своими кодонами на неподвижной 30S субъединице. Получается, что в этот момент обе тРНК окружены чем-то вроде «гибридных» 50S E/30S P-сайтов и 50S P/30S A-сайтов. Наконец, 30S субъединица тоже приходит в движение, так что отставшие антикодоны двух тРНК «догоняют» свои аминоацильные концы. Поскольку гибридное состояние достигается за счет сдвига большой и малой субъединиц по фазе движения относительно мРНК, нормальная структура рибосомы восстанавливается вторым этапом транслокации — выравниванием субъединиц друг относительно друга на мРНК.

Совершая транслокацию, рибосома стоит перед явной дилеммой. Чтобы сдвинуться с места, ей надо порвать множество контактов с тРНК. В то же время ей нужно сохранить сцепление кодонов с антикодонами (даже с деацилированной тРНК нельзя разорвать связь раньше времени). Возможно, совершающая транслокацию рибосома перескакивает между альтернативными (дискретными) конформациями. Скорее всего, «перескок» сопровождается изменениями во вторичной структуре рРНК. Во всяком случае, было установлено, что аккуратность трансляции (точность воспроизведения матрицы) может быть нарушена из-за мутаций в рРНК. Некоторые из этих мутаций влияют на паттерны использования альтернативных комплементарных участков рРНК. Таким образом, точность трансляции может зависеть от того, насколько крепко тРНК удерживается рибосомой при смене альтернативных вариантов вторичной структуры рРНК во время транслокации.

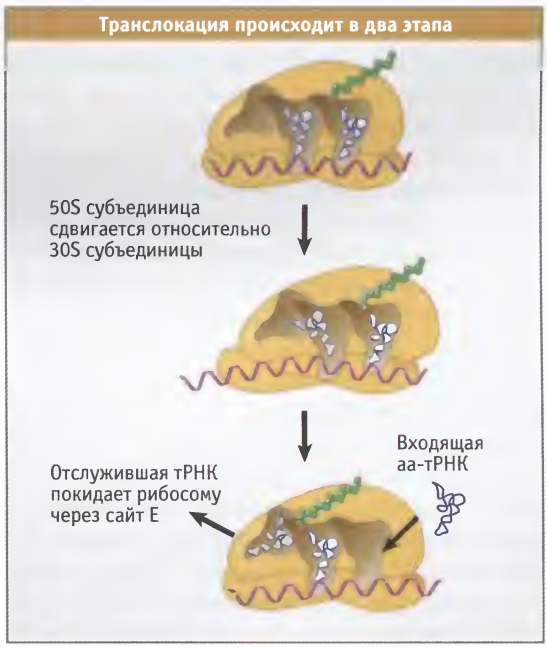


РИС. 8.29 Все существующие модели акта транслокации подразумевают его разбиение на два этапа. Первый этап включает в себя образование пептидной связи; при этом аминоацильный конец тРНК, занимающей сайт А, перемещается в сайт Р. На втором этапе транслокации противоположная часть той же тРНК, несущая антикодон, также перемещается в сайт Р

8.13 Факторы элонгации поочередно связываются с рибосомой

Ключевые понятия:

- Для транслокации нужен фактор EF-G, по своей структуре похожий на комплекс EF-Tu–ГТФ–аминоацил-тРНК.
- Рибосома не может одновременно связывать EF-Tu и EF-G.
- Для транслокации используется гидролиз ГТФ, которым укомплектован EF-G: эта реакция, меняющая конформацию EF-G, ускоряет изменение структуры рибосомы.

При транслокации расходуется ГТФ. Это происходит вследствие работы еще одного фактора элонгации, EF-G. В бактериальной клетке он содержится в значительном количестве: около 20 000 молекул (т. е. ~1 молекула на рибосому).

Рибосомы не умеют связывать EF-Tu и EF-G одновременно. Синтез белка можно уподобить колебательному процессу, изображенному на рис. 8.31: два главных вспомогательных фактора элонгации по очереди заходят в рибосому и покидают ее. Это значит, что EF-Tu–ГДФ уходит до появления EF-G, который должен будет покинуть рибосому до прихода следующего тройственного комплекса EF-Tu–ГТФ–аминоацил-тРНК.

Чем обусловлена несовместимость двух факторов на одной рибосоме? Может быть, какими-то аллостерическими эффектами (если каждый из факторов адаптирует «под себя» конформацию всей рибосомы)? А может быть, их несовместимость обусловлена конкуренцией за перекрывающиеся сайты связывания? На РИС. 8.30 показано необычайное структурное сходство тройственного EF-Tu–ГДФ–аминоацил-тРНК комплекса с фактором EF-G. Своей пространственной структурой EF-G как бы имитирует фактор EF-Tu, связавший акцепторную шпильку аминоацил-тРНК. Глядя на рисунок, сразу хочется предположить, что факторы конкурируют за один общий сайт связывания (где-нибудь недалеко от сайта А). С точки зрения общей организации синтеза белка, необходимость диссоциации одного фактора для связывания другого служит формальной гарантией упорядоченного хода событий.

Оба главных фактора элонгации, EF-Tu и EF-G — это типичные G-белки (небольшие ГТФазы, мономеры). Когда такой белок укомплектован ГТФ, он активен (во многом за счет потенциальной энергии химической связи между фосфатами β и γ в ГТФ). Так, находясь в своих активных формах, EF-Tu и EF-G связываются с рибосомой. В качестве комплектующей части фактора элонгации рибосома признает именно ГТФ; в этом состоит гарантия обеспечения синтеза белка необходимой энергией. При гидролизе ГТФ до ГДФ происходит выброс энергии, и G-белок теряет свою активность. Реакция гидролиза ГТФ катализируется самим G-белком в определенных условиях.

EF-G связывается с рибосомой для катализа транслокации; после транслокации фактор покидает рибосому. Если у EF-G заменить ГТФ очень похожей молекулой, GMP-PCP, которая не может быть гидролизована, транслокация все равно состоится: EF-G как ни в чем не бывало



РИС. 8.30 Структура тройственного комплекса аминоацил-тРНК–EF-Tu–ГТФ (на рисунке слева) аналогична пространственной организации доменов EF-G (справа). Аналогичные домены с консервативной структурой у EF-Tu и EF-G раскрашены, соответственно, в красный и зеленый цвета; тРНК и похожий на нее домен EF-G раскрашены в фиолетовый цвет. Фотография любезно предоставлена Poul Nissen, University of Aarhua, Denmark

свяжется с рибосомой, и она совершит все необходимые движения (правда, гораздо медленнее, чем в норме, из-за отсутствия энергии гидролиза ГТФ). Но вот для того, чтобы фактор EF-G смог покинуть рибосому, необходим гидролиз ГТФ.

Необходимость ухода EF-G для продолжения синтеза белка была доказана с использованием фусидовой кислоты. Этот стероидный антибиотик как бы «замораживает» рибосому в состоянии только что совершенной транслокации (см. РИС. 8.31). В присутствии фусидовой кислоты рибосома может совершить один-единственный акт транслокации: рибосома связывает EF-G, после чего следуют гидролиз ГТФ и продвижение рибосомы на три рибонуклеотидных положения мРНК. Затем, вопреки нормам, EF-G и ГДФ уже не могут покинуть рибосому, потому что фусидовая кислота стабилизирует комплекс рибосомы с EF-G–ГДФ. В результате рибосома не может принять следующую аминоацил-тРНК, и синтез белка останавливается.

Транслокация как функция самой рибосомы обусловлена ее огромной структурной пластичностью (см. разд. 8.17 «В рибосоме есть несколько активных центров»). Активный фактор EF-G нужен для побуждения рибосомы к совершению транслокации. Гидролиз ГТФ происходит непосредственно перед транслокацией; он ускоряет движение субъединиц друг относительно друга. Скорее всего, энергия гидролиза ГТФ находит выражение в изменении конформации EF-G, что и приводит рибосому в движение.

За время транслокации EF-G успевает сильно поменять свою ориентацию относительно рибосомы. Перед

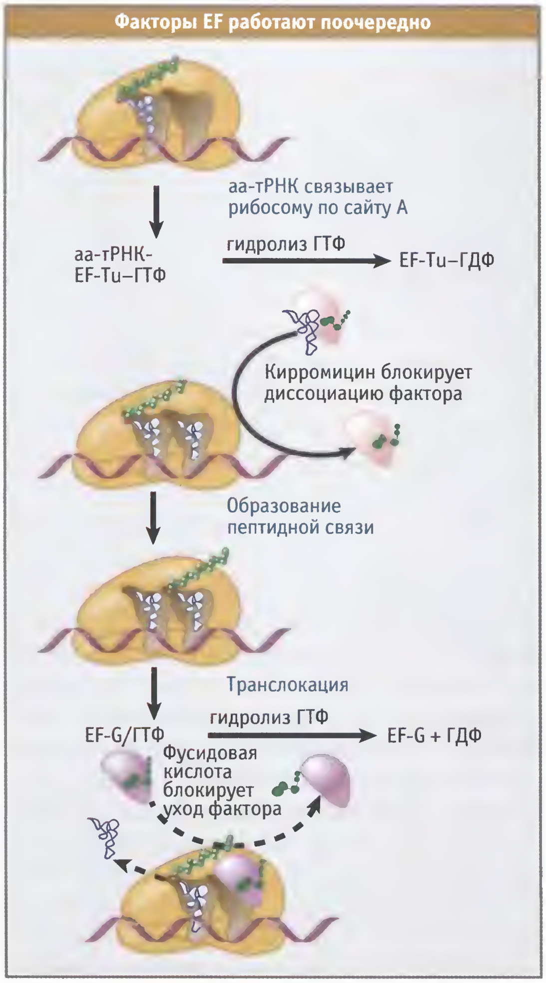


РИС. 8.31 Факторы EF-Tu и EF-G, ведущие элонгацию, поочередно взаимодействуют с рибосомой (по мере того как рибосома принимает новую аминоацил-тРНК, образуется пептидная связь, происходит транслокация, рибосома принимает новую аминоацил-тРНК и т. д.)

транслокацией он расположен поперек двух субъединиц рибосомы. Большинство контактов EF-G с 30S субъединицей образовано его так называемым доменом 4, находящимся в сайте А. Скорее всего, именно этот домен EF-G оттесняет тРНК из сайта А в сайт Р. При завершении транслокации домен 4 фактора EF-G повернут уже к 50S субъединице.

Эукариотическим аналогом EF-G является белок eEF2, который тоже работает как ГТФ-зависимая «транслоказа». Действие eEF2 точно так же ингибируется фусидовой кислотой. Комплекс eEF2 с ГТФ вполне стабилен; его можно выделить из клеток, и он будет связывать рибосомы с последующим гидролизом своего ГТФ *in vitro*.

Дифтерийный токсин (diphtheria toxin) губит эукариотическую клетку путем воздействия на фактор eEF2. Дифтерийный токсин инактивирует eEF2. Используя никотинамидадениндинуклеотид (НАД) в качестве кофактора, токсин пришивает к фактору eEF2 массивный аденозиндифосфорибозил (АДФР). Субстратом служит не совсем обычная аминокислота (продукт модификации гистидина), которая входит в состав eEF2 у многих видов эукариот. Конъюгат АДФР-eEF2 не может участвовать в синтезе белка.

Реакция АДФ-рибозилирования, инактивирующая eEF2, является причиной летального действия дифтерийного токсина. Реакция очень эффективна: чтобы убить клетку, бывает достаточно всего одной молекулы дифтерийного токсина.

8.14 Синтез белка терминируется тремя кодонами

Ключевые понятия:

- Каждый из трех кодонов UAA (охра), UAG (янтарь) и UGA (опал) может терминировать синтез белка.
- Для большинства бактерий, в порядке убывания частоты использования, терминирующие кодоны можно расположить так: UAA>UGA>UAG.

Из четырех рибонуклеотидов A, U, G, C можно составить 64 различных триплета. В соответствии с универсальным генетическим кодом, 61 триплет транслируется в аминокислоты. Три оставшихся триплета служат терминирующими кодонами, или **стоп-кодонами**, которые пресекают рост полипептида на рибосоме. Стоп-кодоны имеют тривиальные названия (названия минералов), уходящие корнями в историю их открытия. Триплет UAG известен как «янтарный», или **амбер-кодон**; UAA — это **охра**; UGA иногда все еще называют **опалом**.

Использование этих трех триплетов в качестве стоп-кодонов было выявлено при проведении генетического теста, позволившего выделить два типа точечных мутаций в произвольной открытой рамке считывания:

- Точечная мутация меняет смысл одного из кодонов: теперь он транслируется по-другому, так что белок синтезируется с одной ошибкой (в одном из положений полипептидной цепочки стоит ошибочная аминокислота). Это **миссенс-мутация**; ее последствия для функции белка зависят от положения и характера замены (насколько похожи две аминокислоты и т. п.).
- Точечная мутация лишает смысла один из кодонов (то есть превращает его в стоп-кодон). Это **нонсенс-мутация**. Она приводит к **преждевременной терминации** синтеза белка: полипептид как бы обрывается, потому что остальная часть открытой рамки считывания, лежащая за стоп-кодоном, теперь не транслируется. Часто это приводит к полной потере функции белка (конечно, все зависит от положения мутации: чем дальше она от начала рамки, тем больше шансов на сохранение функции).

(Иногда к терминирующим триплетам применяют термин нонсенс-кодон, но этот вариант менее корректен, чем термин стоп-кодон.) Во всех генах с из-

вестной нуклеотидной последовательностью первый стоп-кодон следует непосредственно за кодоном, соответствующим С-концевой аминокислоте белка дикого типа (нормального продукта этого гена). Анализ нонсенс-мутаций показывает, что для преждевременной терминации достаточно наличия внутри рамки любого из трех стоп-кодонов. Можно заключить, что наличие триплета UAG, UAA, или UGA является необходимым и достаточным условием завершения синтеза — либо в естественном порядке, то есть в запланированном конце открытой рамки считывания, либо из-за стоп-кодона, порожденного нонсенс-мутацией внутри кодирующей последовательности.

У бактерий наиболее часто используемым терминирующим кодоном является UAA. Из двух других, UGA используется интенсивнее, чем UAG, несмотря на то что UGA — это самый слабый стоп-кодон. Транслируя UGA (опал), рибосома иногда ошибается: интерпретирует его как аминокислоту и продолжает наращивать полипептид до тех пор, пока ей не встретится еще один стоп-кодон.

8.15 Терминирующие кодоны узнаются белковыми факторами

Ключевые понятия:

- Терминирующие кодоны узнаются специальными белками (факторами терминации, или «релизинг-факторами»), а не аминоацил-тРНК.
- Релизинг-факторы класса 1 по структуре напоминают аминоацил-тРНК–EF-Tu и EF-G.
- Релизинг-факторы класса 1 узнают терминирующие кодоны; это узнавание выливается в гидролиз связи между полипептидом и несущей его тРНК.
- Релизинг-факторам класса 1 помогают ГТФ-зависимые релизинг-факторы класса 2.
- У бактерий (обладателей двух разновидностей релизинг-факторов класса 1) и эукариот (с однотипными релизинг-факторами класса 1) механизмы терминации в целом похожи.

Завершение трансляции происходит в две ступени. Сначала происходит собственно *терминация*, то есть снятие полипептидной цепочки с последней тРНК. Следом происходит так называемая *пост-терминация*: отделение тРНК и мРНК от рибосомы и диссоциация самой рибосомы на большую и малую субъединицы.

Терминирующим кодонам не соответствуют никакие тРНК. Они работают принципиально иным способом, за счет специфичного связывания (прямого узнавания) особыми белками — факторами терминации. (Таким образом, этап терминации не обусловлен взаимодействиями типа «кодон–антикодон»; хотя в качестве сигнальных элементов терминации снова выступают рибонуклеотидные триплеты мРНК. Вероятно, последнее обстоятельство отражает ход эволюции генетического кода.)

Терминирующие кодоны узнаются конкретной разновидностью факторов терминации — **релизинг-факторами** (**RF**) класса 1. У *E. coli* таких факторов два, и у этих факторов разная (хотя и перекрывающаяся) специфичность: **RF1** узнает UAA и UAG, **RF2** узнает UAA и UGA. Эффективность узнавания, в частности, зависит от рибонуклеотидного контекста терминирующих кодонов (азотистых оснований, следующих за кодоном в 3′-концевую нетранслируемую область). Оба фактора, RF1 и RF2, оказывают свое действие в сайте А рибосомы (в сайте Р этой рибосомы должна присутствовать полипептидил-тРНК). Содержание релизинг-факторов в каждой клетке ниже содержания факторов инициации и элонгации: всего ~600 молекул RF1 и RF2 на клетку. Это означает, что каждые десять рибосом обслуживаются одной молекулой каждого релизинг-фактора. Скорее всего, релизинг-факторов класса 1 не всегда было два: изначально у бактерий был один-единственный RF, который узнавал все терминирующие кодоны; позднее в эволюции произошла его дивергенция («разделение») на два фактора с разной специфичностью. Это подтверждается тем, что в цитоплазме эукариот есть только релизинг-фактор класса 1, eRF.

Релизинг-факторам класса 1 помогают релизинг-факторы класса 2, у которых нет специфичности ни к каким кодонам (они вообще не взаимодействуют с мРНК). Релизинг-факторы класса 2 укомплектованы молекулами ГТФ. У *E. coli* роль факторов класса 2 всецело относится к высвобождению факторов класса 1 из рибосомы.

Хотя общий принцип терминации у про- и эукариот один и тот же, характер взаимодействия между релизинг-факторами классов 1 и 2 немного отличается.

Бактериальные факторы класса 1, RF1 и RF2, узнают терминирующие кодоны и побуждают рибосому к гидролизу пептидил-тРНК. Реакция отщепления полипептида от тРНК по своему механизму аналогична переносу полипептида (стандартному для элонгации), только вместо аминоацил-тРНК (см. рис. 8.34) акцептором полипептида служит молекула воды.

После отщепления полипептида RF1 (или RF2) покидает рибосому с помощью релизинг-фактора класса 2, **RF3**, который похож на EF-G. Неактивный релизинг-фактор класса 2, RF3-ГДФ, связывает рибосому заранее, еще до того, как возникнет необходимость в терминации, для того чтобы его ГДФ успел замениться на ГТФ (поскольку условия для этой реакции обеспечивает рибосома). Активированный RF3-ГТФ может эффективно взаимодействовать с ГТФазным центром рибосомы; через это взаимодействие RF3 заставляет RF1/2 покинуть рибосому после завершения собственно терминации.

По своей структуре белок RF3 напоминает ГТФ-связывающие домены EF-Tu и EF-G, в то время как RF1 и RF2 похожи на С-концевой домен EF-G, который, в свою очередь, похож на тРНК. Сходство доменов подсказывает, что релизинг-факторы имеют дело с тем же сайтом

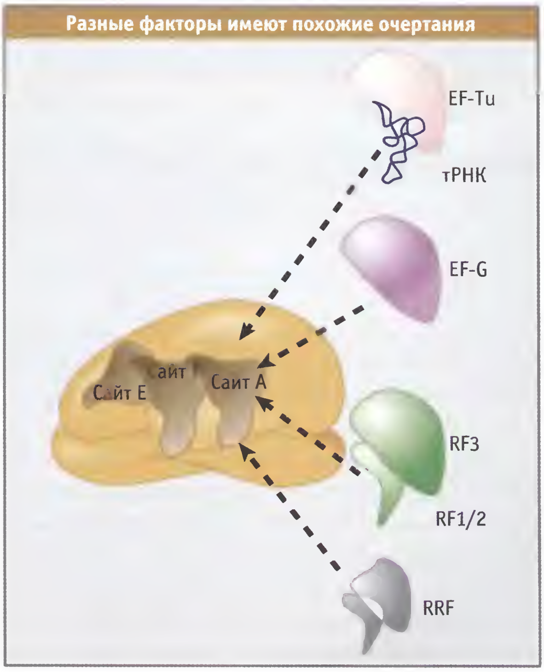


РИС. 8.32 Молекулярная мимикрия позволяет фактору EF-Tu (в комплексе с тРНК), фактору транслокации EF-G, а также факторам терминации, RF1/2-RF3, связываться с одним и тем же сайтом рибосомы. Фактор освобождения рибосомы (RRF, ribosome recycling factor) обсуждается ниже (см. рис. 8.35)



РИС. 8.33 Эукариотический фактор терминации eRF1 имеет доменную организацию, имитирующую очертания и свойства тРНК. Мотив GGQ на верхушке домена 2 играет важную роль в отделении полипептидной цепочки от тРНК. Любезно предоставлено David Barford, The Institute of Cancer Research

рибосомы, что и факторы элонгации. РИС. 8.32 служит иллюстрацией той идеи, что все перечисленные факторы имеют одни и те же очертания и по очереди связываются с одним и тем же сайтом рибосомы (примерно совпадающим с сайтом А или, по крайней мере, значительно перекрывающимся с сайтом А).

Эукариотический релизинг-фактор класса 1, eRF1, может узнать любой из трех терминирующих кодонов. По своей аминокислотной последовательности белок eRF1 совершенно не похож на бактериальные факторы терминации. (Он может терминировать синтез белка без помощи эукариотического фактора класса 2 eRF2 *in vitro*; хотя в некоторых системах, например в дрожжах *in vivo*, терминация без eRF2 не идет.) Структура eRF1 опять же выглядит знакомо: на РИС. 8.33 хорошо различимы три домена, как бы подражающие структуре тРНК.

На верхушке домена 2 расположен очень важный мотив из трех аминокислот, GGQ. По позиции, занимаемой в сайте А, мотив GGQ соответствует аминокислоте (аминоацилу) в составе аминоацил-тРНК. Именно поэтому глутамину (Q из GGQ) так легко удается «обмануть элонгацию», подставив вместо аминоацила молекулу воды (H_2O), на которую и идет перенос полипептида во время терминации. На РИС. 8.34 проведено сравнение реакции терминации с реакцией переноса полипептида, типичной для элонгации. Во время терминации аналогичный перенос идет на гидроксильную группу, взятую

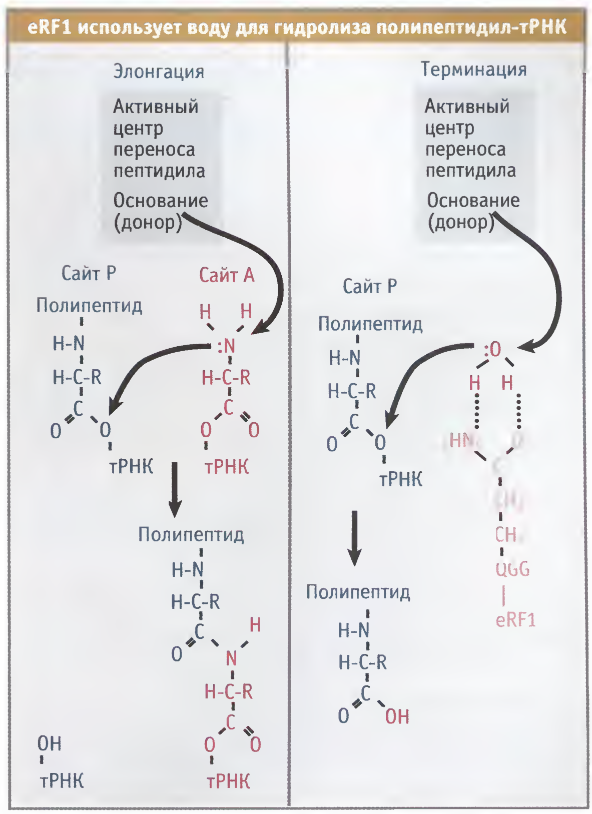


РИС. 8.34 Реакции переноса полипептида (транспептидации) при элонгации и терминация похожи между собой. Каждую из них можно классифицировать как трансэстерификацию. В активном центре фактора, ведущего реакцию, есть основание, которое атакует связь N—H или O—H; при этом N либо O освобождаются для атаки связи полипептида с тРНК

у воды: по сути, это гидролиз связи между завершенным полипептидом и тРНК (см. рис. 8.48 и сопутствующее ему обсуждение работы активного центра пептидилтрансферазы).

Мутации в RF-генах снижают эффективность терминации; это выражается в стремлении рибосом пропускать терминирующие кодоны и читать мРНК дальше, в 3′-концевую область, которая в норме не транслируется. Искусственное усиление уровня экспрессии (сверхэкспрессия) RF1 или RF2 повышает эффективность терминации трансляции соответствующими кодонами. Количественный характер обоих эффектов говорит о том, что в узнавании терминирующих кодонов факторы RF1 и RF2 соперничают с аминоацил-тРНК, пытающимися неверно истолковать терминирующий кодон как один из значимых. На самом деле релизинг-факторы узнают свои терминирующие кодоны очень эффективно.

Как было сказано в начале этого раздела, собственно терминация приводит к отделению готового полипептида, но не освобождает рибосому от деацилированной тРНК и мРНК; к тому же после завершения трансляции неплохо было бы разъединить рибосому на большую и малую субъединицы. На **РИС. 8.35** показано, что для диссоциации компонентов, участвовавших в синтезе белка (тРНК, мРНК, 30S и 50S субъединиц) нужен еще один фактор, RRF (ribosome recycling factor; фактор освобождения рибосомы). RRF действует совместно с EF-G в реакции с расходом энергии гидролиза ГТФ. Подобно другим факторам терминации, RRF своей структурой имитирует тРНК (но в структуре RRF нет эквивалента 3′-концевой аминоацил-связывающей области тРНК).

Последним участником терминации выступает фактор IF-3, уже знакомый нам из описания инициации. Таким образом, круг замкнулся! Теперь синтез белка можно по праву характеризовать как циклический процесс. Фактор IF-3, как мы помним, стабилизирует малые субъединицы рибосомы в диссоциированном состоянии. Во время пост-терминации RRF действует на 50S субъединицу, а действие IF-3 направлено на отрыв деацилированной тРНК от 30S субъединицы. После разделения субъединиц IF-3 так и остается ассоциированным с малой субъединицей; он предотвращает самопроизвольное воссоединение рибосомных субъединиц.

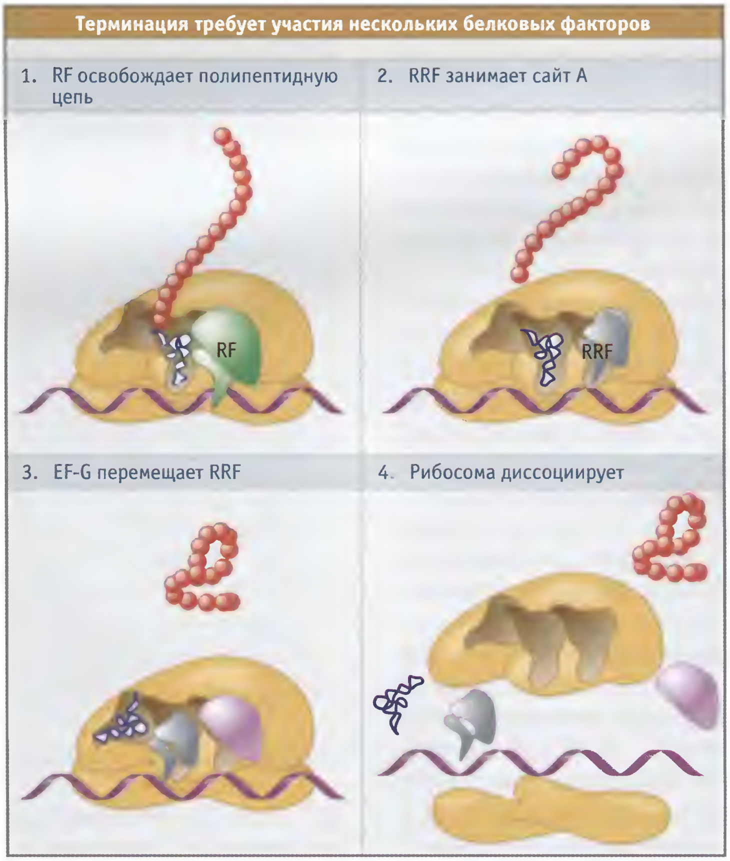


РИС. 8.35 Фактор RF терминирует синтез белка, отделяет полипептид от рибосомы. Фактор RRF освобождает рибосому от оставшейся тРНК, а EF-G отсоединяет RRF, что приводит к диссоциации рибосомы

8.16 Обе субъединицы рибосомы пронизаны рибосомальными РНК

Ключевые понятия:

- Каждая молекула рибосомальной РНК (рРНК) имеет несколько дискретных доменов, формирующих независимые складки.
- Практически все рибосомальные белки находятся в контакте с рРНК.
- Большинство контактов между субъединицами рибосомы — это контакты между 16S и 23S рРНК.

Рибосомальная РНК (рРНК) составляет около двух третей массы бактериальной рибосомы. Наиболее эффективный способ анализа вторичной структуры больших РНК — это сравнение сегментов их последовательности у разных видов организмов. Дело в том, что сегменты, участвующие в образовании вторичной структуры, эволюционно консервативны. Если пара рибонуклеотидов важна для структуры рРНК, она будет находиться в одном и том же положении конкретной рРНК у всех родственных видов. На основании этого подхода (и необходимых контрольных экспериментов) были построены подробные модели 16S и 23S рРНК.

Как выяснилось, каждая большая рРНК состоит из нескольких дискретных доменов. Так, 16S рРНК образует четыре основных домена, в которых только половина рибонуклеотидов взаимодействует между собой по принципу комплементарности (см. рис. 8.45). В 23S рРНК насчитывается шесть основных доменов. В каждом домене есть множество двухцепочечных участков, в отдельности довольно коротких (в среднем менее 8 пн) и зачастую несовершенных (с выпячиваниями неспаренных рибонуклеотидов). Похожие модели получены и для рРНК митохондрий (они короче бактериальных, а доменов в них меньше), и для цитоплазматических рРНК эукариот (они длиннее, доменов в них больше). Увеличение длины эукариотических рРНК происходит главным образом за счет добавления целых доменов. Анализ кристаллической структуры рибосом показал, что внутри больших и малых субъединиц домены рРНК складываются независимо, причем каждый домен занимает в субъединице рибосомы свое особое положение.

16S рРНК в составе 30S субъединиц, в составе целых 70S рибосом, а также в составе целых рибосом, ведущих синтез белка, проявляет разную реакционную способность к воздействию некоторых химических агентов. Изменения в реакционной способности 16S рРНК происходят в момент связывания мРНК малой субъединицей, в момент ассоциации большой и малой субъединиц, а также в момент связывания тРНК. Некоторые из этих изменений являются прямым следствием взаимодействия рРНК с мРНК или тРНК; другие изменения (в реакционной способности рРНК) сопряжены с изменениями в структуре рибосомы. Оба типа изменений свидетельствуют о высокой конформационной динамичности рибосомы, ведущей синтез белка.

У рРНК, в ее первичной структуре, есть одна характерная особенность: наличие в определенных положениях метильных групп. В 16S рРНК насчитывается ~10 метильных групп (в основном находящихся вблизи 3′-конца молекулы). В 23S рРНК их насчитывается около 20. В клетках млекопитающих большие рРНК, 18S и 28S, несут еще больше метильных групп (соответственно 43 и 74). Эти цифры означают, что в больших рРНК млекопитающих метилировано ~2% рибонуклеотидов (примерно втрое больше, чем у бактерий).

Большие субъединицы всех рибосом (за исключением митохондриальных) также содержат еще по одной малой рРНК. У бактерий это **5S рРНК** длиной 120 оснований. При том, что ее последовательность менее консервативна, чем последовательности больших рРНК, все 5S рРНК имеют крайне высокую степень внутренней комплементарности.

В цитоплазматических рибосомах эукариот есть своя малая РНК — **5,8S рРНК**. По своей последовательности (первичной структуре) она соответствует 5′-концевой области бактериальной 23S рРНК. (Заметим, что в дополнение к упомянутым, в состав большой субъединицы эукариотической рибосомы входит еще одна РНК, 5S рРНК.)

Что касается рибосомальных белков, некоторые из них проявляют сильное сродство к свободной рРНК. Другие, напротив, начинают взаимодействовать с рРНК только в связанной с белком форме. Это говорит о важности конформации рРНК для возникновения сайтов связывания некоторых белков. По мере поступления каждый белок индуцирует в рРНК определенные конформационные изменения, облегчающие присоединение других рибосомальных белков (с определенной очередностью). У *E. coli* практически каждый белок из состава малой (30S) субъединицы напрямую взаимодействует с 16S рРНК, хотя прочность этих взаимодействий (сила, число межмолекулярных контактов и т. п.) для разных белков отличается. Строение сайтов связывания рРНК в рибосомальных белках весьма разнообразно, и это указывает на то, что, скорее всего, рибосомальные белки узнают рРНК и связывают ее по разным молекулярно-кинетическим «сценариям».

Полная 70S рибосома имеет асимметричное строение. На **рис. 8.36** дана морфологическая схема 30S субъединицы, в которой хорошо различимы четыре отдела: головка, шея, тело и платформа. На **рис. 8.37** дана аналогичная схема 50S субъединицы, которую легко узнать по центральному протуберанцу (central protuberance), в котором находится 5S рРНК, и основному массиву (stalk), где находится множество копий белка L7. На **рис. 8.38** показано, что основание малой субъединицы входит в выемку (зарубку; notch) на большой субъединице; при этом между субъедини-

![fig](


РИС. 8.36 В 30S субъединице есть головка, отделенная шеей от тела с плоским выступом-основанием



РИС. 8.37 У 50S субъединицы есть центральный протуберанец (где находится 5S рРНК), выемка и основной массив, в котором доминируют копии белка L7



РИС. 8.38 Платформа 30S субъединицы входит в выемку на 50S субъединице: образуется 70S рибосома

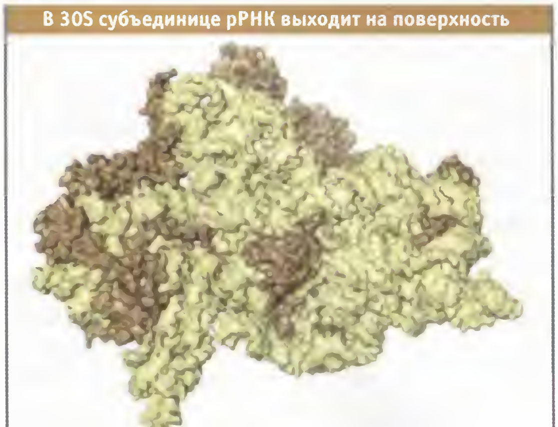


РИС. 8.39 30S субъединица рибосомы представляет собой рибонуклеопротеиновую частицу. Белки раскрашены в желтый цвет. Любезно предоставлено V. Ramakrishnan, Medical Research Council (UK)

цами образуется полость, в которую открывается ряд важных сайтов.

Про малую субъединицу можно сказать, что вся ее структура следует пространственной организации 16S рРНК (то есть каждый морфологический отдел субъединицы соответствует какому-нибудь домену рРНК). Тело строится на 5′-концевом домене, основание — на центральном домене, а головка — на 3′-концевом домене. На РИС. 8.39 показано, как асимметрично распределены белок и РНК в малой субъединице рибосомы. Основание (платформа), находящееся на стыке субъединиц, почти полностью состоит из РНК. Из множества r-белков малой субъединицы в формировании основания задействовано всего два: небольшой фрагмент S7 и, скорее всего, часть S12. Таким образом, ассоциация и диссоциация субъединиц должны зависеть от контактов между компонентами большой субъединицы и 16S рРНК. Ассоциация субъединиц может быть нарушена мутацией в одноцепочечной петле 16S рРНК (в положении 791); эта петля лежит как раз в плоскости соприкосновения двух субъединиц.

Как показали эксперименты по внесению серийных модификаций, рибонуклеотиды в некоторых других положениях 16S рРНК также имеют отношение к ассоциации субъединиц. Можно утверждать, что в рибосоме именно рРНК является безусловной структурной и функциональной доминантой, это согласуется с идеей об эволюционно предковой форме рибосомы как о частице, в основном состоящей из РНК, а не из белка.

Большая (50S) субъединица бактериальной рибосомы характеризуется более однородным распределением компонентов, чем малая субъединица. Ее структура опирается на массив из взаимно-пересекающихся коротких элементов вторичной структуры рРНК (по сути, это плотная укладка из коротких спиралей). Наружная поверхность рибосомы почти полностью замаскирована белком; исключение представляет активный центр пептидил-трансферазы (см. разд. 8.19 «23S рРНК обладает пептидил-трансферазной активностью»). Что касается 23S рРНК, почти все ее сегменты взаимодействуют с рибосомальными белками; правда, многие из этих белков довольно аморфны.

Соединение большой и малой субъединиц в 70S рибосоме можно интерпретировать в терминах контактов между 16S рРНК и 23S рРНК. Вспомогательными являются контакты рРНК одной субъединицы с белками другой субъединицы, а также немногочисленные белок-белковые контакты двух субъединиц. На схеме вторичных структур отдельных доменов больших рРНК РИС. 8.40 отмечены точки контактов с рРНК другой субъединицы рибосомы. РИС. 8.41 как бы раскрывает рибосому: на рисунке 30S и 50S субъединицы развернуты друг от друга относительно вертикальной оси. Такой ракурс позволяет увидеть местонахождение точек контакта на поверхностях соприкосновения двух субъединиц.



РИС. 8.40 Точки контакта между рРНК сосредоточены в двух доменах 16S рРНК и одном домене 23S рРНК. Воспроизведено по: Yusupov, M. M. 2001. Science. 292: 883–896. © 2001 AAAS

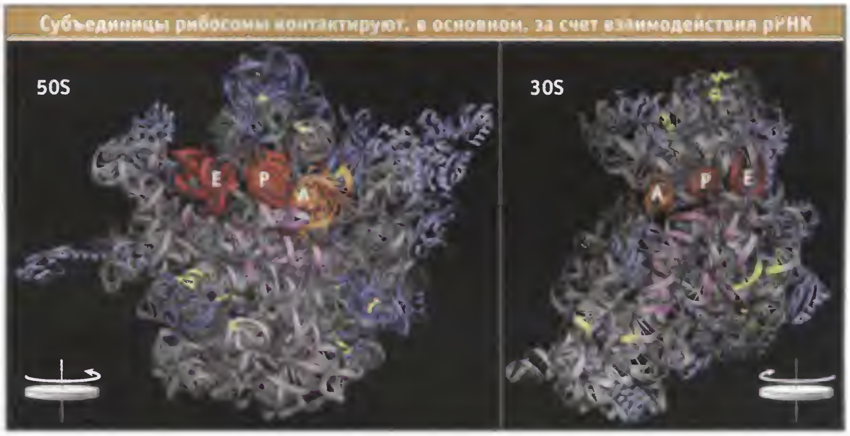


РИС. 8.41 Контакты между субъединицами рибосомы — это главным образом контакты между их рРНК (на модели они раскрашены в фиолетовый цвет). Контакты, образованные при участии белков, обозначены желтым цветом. Субъединицы специально развернуты друг от друга, чтобы были видны те части их поверхности, где находятся точки контакта: 50S субъединица повернута на 90° против часовой стрелки, а 30S субъединица — на 90° по часовой стрелке (противоположно привычной ориентации). Любезно предоставлено Harry Noller, University of California, Santa Cruz

8.17 В рибосоме есть несколько активных центров

Ключевые понятия:

- Взаимодействия с участием рРНК служат главной движущей силой в работе рибосомы.
- Облик тРНК-связывающих сайтов в значительной степени определяется свойствами конкретных участков рРНК.

Чтобы хоть как-то понять принцип функционирования рибосомы, нужно запомнить одну важную вещь: *рибосома является кооперативной структурой*, всецело зависящей от взаимодействия своих активных сайтов. Эти сайты не так малы, как активные центры большинства ферментов. Напротив, это довольно обширные функциональные отделы рибосомы; их организация и режим работы определяются локальными свойствами рРНК (во всяком случае, не в меньшей степени, чем рибосомальными белками). Анализ кристаллических структур субъединиц бактериальной рибосомы, вместе и по отдельности, позволяет довольно хорошо понять и общую организацию активных сайтов рибосомы, и определяющую роль рРНК на каждом этапе синтеза белка (и даже на каждом минимальном шаге каждого этапа). Самые свежие данные из этой серии, кристаллическая структура с разрешением 5,5 Å позволяет четко детектировать местоположение молекул тРНК и функциональных сайтов. На сегодняшний день мы уже можем судить о функциях рибосомы с позиций ее структуры.

Если пренебречь тонкостями механизмов инициации, элонгации и терминации, работу рибосомы можно

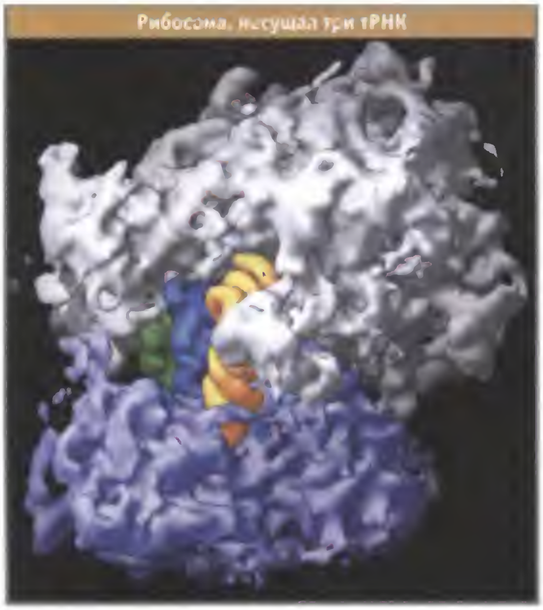


РИС. 8.42 70S рибосома, состоящая из белой 50S субъединицы и фиолетовой 30S субъединицы, несет три тРНК, расположенных на ее поверхности: желтую тРНК в сайте А, синюю тРНК в сайте Р, зеленую тРНК в сайте Е. Любезно предоставлено Harry Noller, University of California, Santa Cruz

свести к взаимодействиям с тРНК. На РИС. 8.42 изображена бактериальная 70S рибосома и отмечены положения молекул тРНК в трех сайтах этой рибосомы. Молекулы тРНК в сайтах А и Р расположены почти параллельно друг другу. Все три тРНК выровнены относительно рибосомы: их антикодоновые петли взаимодействуют с мРНК в бороздке 30S субъединицы, а остальные части взаимодействуют с 50S субъединицей. Окружение трех

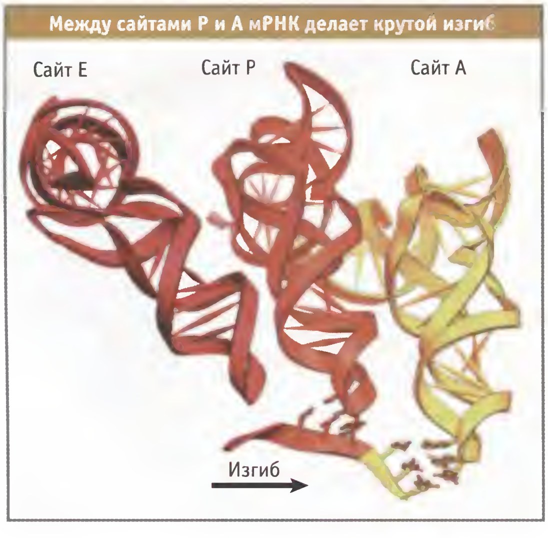


РИС. 8.43 Три тРНК, одновременно находящиеся в рибосоме, имеют разную ориентацию. Изгиб мРНК между сайтами А и Р позволяет двум аминоацил-тРНК связывать смежные кодоны. Любезно предоставлено Harry Noller, University of California, Santa Cruz

тРНК состоит, главным образом, из рРНК. В каждом из трех сайтов рРНК образует с тРНК характерный консервативный паттерн контактов.

При взгляде на эту картину озадачивает то обстоятельство, что две громоздких тРНК умещаются на двух смежных кодонах. Очевидно (и это подтверждается данными рентгеноструктурного анализа), что между сайтами Р и А мРНК делает крутой изгиб (45°), так что две тРНК не соперничают друг с другом из-за места; они как бы смотрят в разные стороны (как показано на РИС. 8.43). Две молекулы тРНК образуют друг с другом угол в 26° (вершина угла приходится на антикодоны). Их 3′-концы предельно сближены — до расстояния в 5 Å (перпендикуляного плоскости рисунка). Такое сближение создает возможность для переноса полипептидной цепочки с пептидил-тРНК в сайте Р на аминоацил-тРНК в сайте А.

Аминоацил-тРНК приносится в сайт А фактором EF-Tu; ее соответствие антикодону является необходимым условием гидролиза ГТФ и освобождения EF-Tu из рибосомы (см. разд. 8.10 «Фактор элонгации EF-Tu загружает аминоацил тРНК в сайт А»). Сначала EF-Tu ставит аминоацил-тРНК на малую субъединицу рибосомы, где антикодон образует пару с кодоном. Чтобы полностью войти в сайт А, тРНК совершает еще одно движение, от которого ее 3′-конец попадает в активный (каталитический) центр пептидил-трансферазы на большой субъединице. Как именно выглядит это движение, никто не знает; есть разные модели. Возможно, тРНК образует L-образный изгиб (колено) между плечами D и TψC, и этим коленом она как бы въезжает в рибосому, позволяя плечу TψC вступить в контакт с рРНК. Согласно другой модели, антикодоновая петля, взаимодействуя с кодоном, становится центром вращения всей тРНК: за счет внутренних изменений остальная часть молекулы тРНК переваливается с одной стороны антикодоновой петли (3′) на другую (5′). В любом случае, независимо от механизма движения, после того как оно произошло, EF-Tu гидролизует свой ГТФ, и синтез белка продолжается.

Значительные изменения положения тРНК внутри рибосомы происходят при транслокации. Антикодоны тРНК совершают перемещения из сайта А в сайт Р на ~28 Å и из сайта Р в сайт Е на ~20 Å. Из-за отклонения от вертикали основная масса каждой молекулы тРНК преодолевает гораздо большее расстояние: 40 Å из сайта А в сайт Р и 55 Å из сайта Р в сайт Е. Должно быть, такие перемещения сопряжены с большими деформациями всей структуры рибосомы во время транслокации.

Многие годы считалось, что транслокация может идти только в присутствии фактора EF-G. Но оказалось, что антибиотик спарзомицин (sparsomycin), являясь ингибитором пептидил-трансферазной активности, тем не менее запускает транслокацию. Спарзомицин ингибирует активность пептидил-трансферазы (это собственная активность рибосомы) за счет блокирования взаимодействия между пептидил- и аминоацил-тРНК. Скорее всего, взаимодействуя со спарзомицином, пептидил-тРНК обретает конформацию, характерную для пост-транслокации, и, благодаря этому, приходит в движение. Во всяком случае,

приходится предположить, что энергия для проведения транслокации в норме берется из самой рибосомы после того, как в ней состоялось образование пептидной связи. EF-G, действуя на рибосому, лишь заставляет ее быстро «сбросить» эту энергию в форме транслокации (точно такое же действие оказывает спарзомицин). Таким образом (и это очень важно), способность к транслокации была признана свойством самой рибосомы.

Модель гибридного состояния гласит, что транслокация происходит в два этапа. Сначала субъединицы рибосомы сдвигаются друг относительно друга; в этом промежуточном состоянии у рибосомы есть два гибридных сайта связывания тРНК: 50SE/30S P и 50SP/30S A (см. рис. 8.29). Интересно, что в первичной структуре рРНК участки, обеспечивающие ассоциацию субъединиц буквально соседствуют с участками, ответственными за взаимодействие с тРНК. По-видимому, сайты связывания тРНК примыкают к поверхности соприкосновения субъединиц рибосомы так, что смещение субъединиц друг относительно друга может быть механически сопряжено с движением тРНК.

На самом деле, рибосома в основном-то и состоит из активных сайтов. Схема расположения этих сайтов, занимающих около двух третей рибосомы, представлена на РИС. 8.44. тРНК входит в сайт А, при транслокации переходит в сайт Р, затем покидает (бактериальную) рибосому через сайт Е. Сайты А и Р простираются на обе субъединицы рибосомы, причем взаимодействие тРНК с мРНК происходит в пределах малой субъединицы, а перенос полипептида — в пределах большой субъединицы. Сайты А и Р примыкают друг к другу; это важно для перемещения тРНК из А в Р-сайт при транслокации. Сайт Е находится рядом с Р; сайт Е — это промежуточная «стоянка» тРНК на пути из рибосомы. Активный (каталитический) центр пептидилтрансферазы находится в пределах 50S субъединицы, в непосредственной близости от аминоацильных концов тРНК в сайтах А и Р (см. разд. 8.18 «16S рРНК играет активную роль в синтезе белка»).

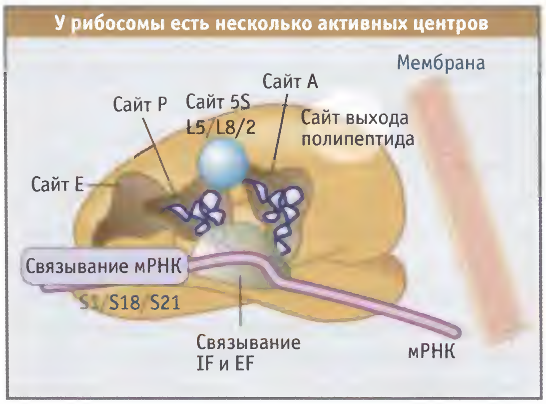


РИС. 8.44 У рибосомы есть несколько активных центров. Рибосома может быть ассоциирована с мембраной. Проходя через сайты А и Р, мРНК делает крутой изгиб. Сайт Е (в данной проекции) находится позади сайта Р. Необозначенный сайт пептидилтрансферазы протянут через верхушки сайтов А и Р. В основании сайтов А и Р находится часть сайта связывания EF-Tu/G

Все ГТФ-связывающие факторы, принимающие участие в синтезе белка (EF-Tu, EF-G, IF-2, а также RF-1, -2, -3) садятся в один и тот же сайт рибосомы, который можно назвать ГТФазным центром, и скорее всего, этот центр содействует запуску гидролиза ГТФ. ГТФазный центр находится в основном массиве большой субъединицы, где доминируют белки L7 и L12. (Белок L7 — это модифицированный L12, с ацетильной группой на N-конце.) Еще один r-белок, L11, в комплексе с участком 23S рРНК длиной 58 осн., образует сайт связывания некоторых антибиотиков, нарушающих ГТФазную активность факторов. Тем не менее ни один из сайтов рибосомы не обладает собственной ГТФазной активностью; роль рибосомы в работе G-факторов заключается в запуске гидролиза ГТФ в случае правильного связывания факторами своих сайтов.

Для первичного связывания 30S субъединицы с мРНК нужен белок S1, у которого есть сильное сродство к нуклеиновым кислотам в одноцепочечной форме. Белок S1 отвечает за денатурацию вторичной структуры (поддержание одноцепочечного состояния) мРНК, связавшейся с 30S субъединицей. Действие S1 необходимо для того, чтобы мРНК пребывала в расплетенном состоянии, пригодном для трансляции. S1 имеет вытянутую форму и взаимодействует с белками S18 и S21. Втроем они образуют комплексный домен, участвующий в первичном связывании малой субъединицы с мРНК и в связывании инициаторной тРНК с комплексом инициации. мРНК взаимодействует с малой субъединицей в месте углубления «основания» (см. рис. 8.36); именно сюда выходит 3'-концевая область 16S рРНК, сегмент которой комплементарен последовательности Шайна–Дальгарно в сайте инициации мРНК.

Факторы инициации тоже связывают эту область рибосомы. Фактор IF-3 можно химически зафиксировать на взаимодействии с 3'-концом рРНК, равно как и с несколькими r-белками, часть которых, судя по всему, взаимодействует с мРНК. Возможно, IF-3 стабилизирует связывание мРНК с 30S субъединицей. При вступлении 50S субъединицы в инициаторный комплекс IF-3 покидает рибосому.

В условиях *in vitro* включение 5S рРНК в состав 50S субъединицы зависит от способности определенного числа молекул белков L5, L8 и L25 образовывать комплекс с этой РНК. Комплекс может связывать 23S рРНК (при том, что ни один из его компонентов по отдельности, не может). В большой субъединице рибосомы этот комплекс занимает место неподалеку от сайтов Р и А.

Синтезируемый полипептид проходит сквозь рибосому, в направлении от активных сайтов к поверхности большой субъединицы — к участку прикрепления к мембране (см. гл. 10 «Локализация белка»). Для выхода полипептидной цепи в рибосоме предусмотрен специальный туннель, ведущий от активного центра пептидил-трансферазы к поверхности большой субъединицы. Этот туннель

узкий (1–2 нм в ширину) и довольно длинный (~10 нм); его стенки состоят в основном из рРНК. Расстояние между пептидил-трансферазным центром и точкой выхода растущего полипептида на поверхность рибосомы составляет ~15 Å. В туннеле помещается около 50 аминокислот; скорее всего, туннель препятствует фолдингу находящегося в нем участка полипептидной цепи.

8.18 16S рРНК играет активную роль в синтезе белка

Ключевые понятия:

- 16S рРНК играет активную роль в функционировании 30S субъединицы рибосомы. 16S рРНК напрямую взаимодействует с мРНК, 50S субъединицей, а также с антикодонами тРНК в сайтах P и A.

Рибосома традиционно рассматривалась как комплекс белков с различными энзиматическими активностями, удерживаемых вместе за счет белок-белковых взаимодействий и связи с рРНК. Открытие собственной каталитической активности некоторых молекул РНК (см. гл. 26 «Сплайсинг и процессинг РНК») заставило задуматься о возможно более активной роли рРНК в работе рибосомы. Были получены экспериментальные подтверждения взаимодействия рРНК с тРНК и мРНК для всех этапов трансляции и доказана необходимость рибосомальных белков для придания рРНК активной конформации. Как выяснилось, у некоторых участков рРНК есть вполне конкретные функции:

- 3′-концевая область 16S рРНК напрямую взаимодействует с мРНК во время инициации.
- Определенные участки 16S рРНК напрямую взаимодействуют с антикодон-содержащими участками тРНК — и в сайте A, и в сайте P. Сходным образом, 23S рРНК взаимодействует с акцепторной шпилькой (CCA-концом) пептидил-тРНК в сайтах P и A.
- Взаимодействие между субъединицами рибосомы относится во многом на счет контактов между 23S и 16S рРНК (см. разд. 8.16 «Обе субъединицы рибосомы пронизаны рибосомальными РНК»).

Большой объем информации о деталях синтеза белка у бактерий был получен в экспериментах с использованием антибиотиков, блокирующих этот процесс на определенных этапах. Мишень любого антибиотика идентифицируется мутациями, придающими бактерии устойчивость к этому антибиотику. Кроме антибиотиков, действующих на рибосомальные белки, было открыто несколько антибиотиков, мишенью которых служит непосредственно рРНК. Это открытие стало еще одним подверждением активной роли рРНК во многих (если не во всех) аспектах работы рибосомы.

К исследованию функций рРНК можно подойти по-разному. Один из подходов заключается в отождествлении конкретных участков рРНК с важными сайтами рибосомы (например, при помощи внесения в рРНК локализованных химических модификаций, влияющих на функции рибосомы) — это так называемый структурологический подход. Другой подход, молекулярно-генетический, заключается в картировании мутаций рРНК и анализе их разнообразных фенотипических эффектов. Карта функциональных сайтов 16S рРНК, построенная по результатам применения двух подходов, дана на РИС. 8.45.

Особая функциональная значимость 3′-концевой области 16S рРНК подтверждается повышенной чувствительностью этой области к смертельному (для *E. coli*) яду колицину. Этот агент (colicin E3), продуцируемый некоторыми бактериями, отщепляет от 3′-конца 16S рРНК *E. coli* фрагмент в ~50 рибонуклеотидов. Своим действием колицин полностью прекращает инициацию синтеза белка в клетке *E. coli*, поскольку от него страдает сразу несколько важных функций: связывание фактора IF-3, узнавание мРНК, связывание тРНК.

3′-концевая область 16S рРНК является непосредственным участником инициации трансляции, в частности, в силу того, что один из ее сегментов образует дуплекс с последовательностью Шайна–Дальгарно сайта связывания рибосомы в составе мРНК (см. рис. 8.16). Другую прямую обязанность 3′-концевой области 16S рРНК можно иллюстрировать фенотипом нуль-мутантов по гену *ksg*. Такие мутанты проявляют устойчивость к антибиотику касугамицину, блокирующему инициацию синтеза белка. У устойчивых к нему мутантов *ksgA* нет фермента метилазы, который в норме навешивает четыре метильных группы на два смежных аденина в конкретном сайте 16S рРНК, неподалеку от ее 3′-конца. Эта реакция порождает высококонсервативный паттерн G–m_2^6A–m_2^6A, обнаруживаемый и у про-, и у эукариот. Этот сайт метилирования играет важную роль в правильной ассоциации 30S субъединицы с 50S субъединицей, которая, в свою очередь, обеспечивает удержание инициаторной тРНК на образовавшейся полной рибосоме. В присутствии касугамицина нормальная (метилированная) рибосома теряет fMet-$тРНК_f$, а рибосома *ksgA*, напротив, способна удерживать инициаторную тРНК даже в присутствии касугамицина.

В продолжение синтеза белка структура 16S рРНК претерпевает множество изменений. Эти изменения можно отслеживать по паттернам защиты рибонуклеотидных положений от разрушения химическими агентами. Защищенные сайты 16S рРНК можно поделить на несколько групп, собранных в малом 3′-концевом и центральном доменах. Хотя в линейной рибонуклеотидной последовательности 16S рРНК защищенные сайты выглядят достаточно рассеянно, в пространственной структуре защищенные сайты с общей функцией явно сближены.

Некоторые из комплексных изменений в паттерне защиты 16S рРНК запускаются конкретными событиями: присоединением 50S субъединицы, связыванием с мРНК, связыванием тРНК. Эти события сопряжены с конформационными изменениями, влияющими на досягаемость рРНК (с поверхности рибосомы). На

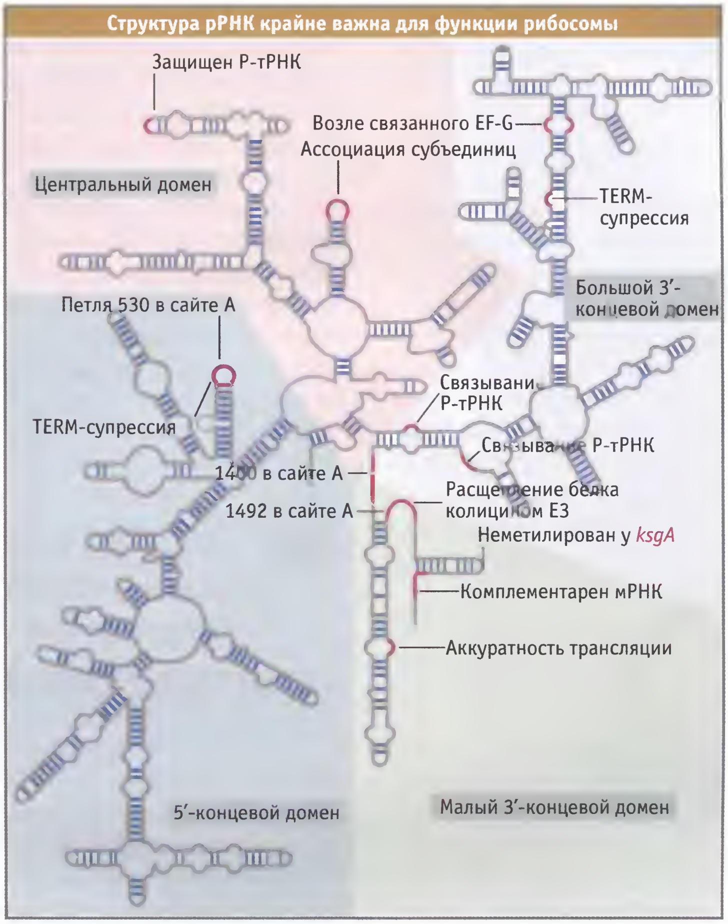


РИС. 8.45 Некоторые сайты 16S рРНК защищены от мечения химическими группами, если 50S субъединица ассоциирована с 30S субъединицей или когда аминоацил-тРНК связывается с рибосомой в сайте А. На другие сайты приходятся мутации, нарушающие ход синтеза белка. Мутации по сайтам TERM-супрессии могут препятствовать терминации трансляции. Большие цветные клетки соответствуют четырем доменам 16S рРНК

РИС. 8.46 приведен частный случай изменения вторичной структуры 16S рРНК: небольшой локальный сдвиг меняет свойства двухцепочечного участка. Изменения в паттерне защиты, однако, ничего не говорят о непосредственном участии рРНК в событиях, с которыми ассоциированы эти изменения.

16S рРНК играет важную роль в работе активных сайтов Р и А. Когда эти сайты не пусты (активны), структура 16S рРНК претерпевает значительные изменения. Некоторые участки попадают под защиту тРНК, связавшейся в сайте А (см. рис. 8.45). Один из них — петля 530, на которую приходятся мутации, вынуждающие рибосому пропускать терминирующие кодоны UAA, UAG и UGA (то есть эти мутации мешают терминации). Другой — это 1400–1500 регион 16S рРНК, закрываемый тРНК в сайте А (называемый так потому, что между положением 1399 и двумя аденинами 1492–1493 содержит два одноцепочечных сегмента, соединенных длинной шпилькой). Интересно, что любой эффект, который оказывает на 16S рРНК связывание с тРНК, можно воспроизвести, взяв вместо целой тРНК лишь олигонуклеотид, соответствующий изолированной антикодоновой шпильке. Таким образом, тРНК-30S взаимодействия в основном затрагивают именно этот регион.

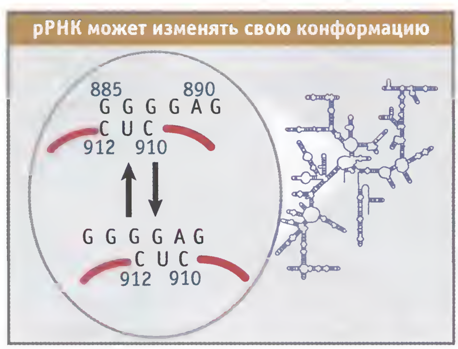


РИС. 8.46 В ходе синтеза белка конформация 16S рРНК может меняться

Аденины в смежных положениях 1492 и 1493 обеспечивают контроль точности образования кодон-ан-

тикодонового комплекса. Оказывается, степень совершенства первых двух рибонуклеотидных пар в дуплексе кодон-антикодон находит отображение в структуре 16S рРНК, причем аденины в положениях 1492 и 1493 служат чем-то вроде щупа в малой бороздке этого дуплекса. Модификация по атому N1 в любом из этих двух аденинов мешает тРНК прочно занять сайт А; такой же эффект оказывают и точечные замены этих аденинов на другие азотистые основания. Эффекты таких мутаций могут быть компенсированы фторированием атомов 2′ в рибонуклеотидах мРНК. На РИС. 8.47 показано, что спаривание кодона с антикодоном позволяет атому N1 каждого аденина взаимодействовать с 2′-ОН в остове мРНК. Если в сайт А входит ошибочная тРНК, структура дуплекса будет далека от совершенства, и указанного воздействия не возникнет. Если же дуплекс получился правильным (то есть если пришедший антикодон подошел к кодону, как ключ к замку), аденины 1492 и 1493 16S рРНК стабилизируют тРНК в сайте А.

Аналогичным образом, тРНК, занимающая сайт Р, закрывает в 16S рРНК множество рибонуклеотидных положений. Хотя в последовательности рРНК эти положения разнесены далеко друг от друга, в пространственной структуре они сцеплены. На самом деле, тРНК в сайте Р имеет больше контактов с 16S рРНК, чем тРНК в сайте А. Возможно, именно этим обусловлена повышенная стабильность связывания пептидил-тРНК по сравнению с аминоацил-тРНК. В этом есть своя логика: если уж тРНК достигла сайта Р, значит, рибосома уже давно признала ее связывание корректным, тогда как в сайте А вопрос о связывании только решается. Регион немного выше 1400 может быть химически «пришит« (crosslinked) к пептидил-тРНК; это дает основание предполагать, что эта область 16S рРНК является структурным компонентом сайта Р.

Основной вывод из изложенных результатов звучит так: рибосомальная РНК образует множество контактов с тРНК и мРНК, причем эти контакты возникают заново в каждом рабочем цикле рибосомы.

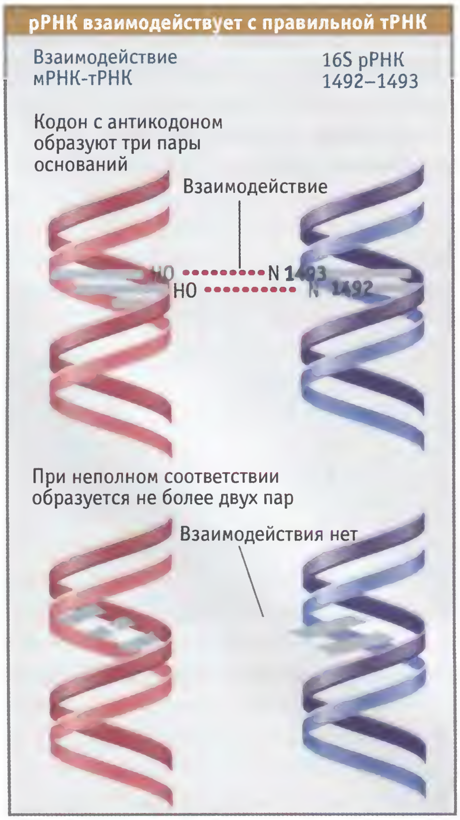


РИС. 8.47 Полное соответствие между кодоном и антикодоном позволяет тРНК взаимодействовать с аденинами 1492-1493 16S рРНК. Если соответствие неполное, взаимодействия нет

8.19 23S рРНК обладает пептидилтрансферазной активностью

Ключевые понятия:

- Пептидилтрансферазная активность, присущая рибосоме, обеспечивается исключительно в 23S рРНК.

Сайты 23S рРНК, обеспечивающие ее функциональную активность, изучены не так хорошо, как для 16S рРНК. Но, по крайней мере, ясно, что для двух больших рРНК действует один и тот же общий принцип отображения структуры в функцию: рибонуклеотиды в определенных положениях обеспечивают выполнение специфичных функций. Устойчивость некоторых рибонуклеотидных положений 23S рРНК к химической деградации зависит от текущей конформации сайтов А и Р. В частности, олигонуклеотиды, совпадающие с акцепторным ССА-концом тРНК, закрывают в 23S рРНК тот же набор положений, что и вся пептидил-тРНК. Следовательно, основные взаимодействия 23S рРНК с пептидил-тРНК в сайте Р приходятся на 3′-концевую область тРНК; аналогично контакты 16S рРНК с тРНК приходятся на антикодоновую шпильку (см. разд. 8.18 «16S рРНК играет активную роль в синтезе белка»).

тРНК контактирует с 23S рРНК и в сайте Р, и в сайте А. В сайте Р гуанин в 2552 положении 23S рРНК образует пару с комплементарным цитозином в позиции 74 пептидил-тРНК. Соответственно, мутация с заменой гуанина 2552 другим азотистым основанием мешает 23S рРНК взаимодействовать с тРНК, однако этот эффект может быть нейтрализован компенсирующей мутацией с заменой цитозина 74 на акцепторном конце тРНК другим азотистым основанием. Тем временем в сайте А смежный гуанин 23S рРНК, G2553, образует пару с С75 аминоацил-тРНК. Таким образом, можно отметить вклад рРНК в оба тРНК-связывающих сайта. Когда мы будем иметь четкое представление этой области рибосомы, перемещение тРНК из сайта А в сайт Р можно будет описать в терминах разрыва и установления контактов с рРНК.

На бактериальной рибосоме есть еще один сайт связывания тРНК, сайт E, который почти не выходит за пределы большой субъединицы; в 23S рРНК идентифицирован ряд рибонуклеотидных положений, защищенность которых определяется конформацией сайта E.

До сих пор мы оставляли открытым очень важный вопрос: какова же природа пептидил-трансферазной активности, которой, как нам известно, обладает сама рибосома, а точнее, большая субъединица рибосомы? О несомненном участии 23S рРНК свидетельствует наличие в ней области, в которой «гнездятся» мутации, придающие клетке устойчивость к антибиотикам — ингибиторам пептидилтрансферазы.

Долгие поиски рибосомальных белков с каталитической активностью не увенчались успехом. Напротив, все данные говорят о том, что каталитическую активность следует приписать рибосомальной РНК большой субъединицы. После удаления с большой субъединицы почти всех белков (на 23S рРНК остается менее 5% исходной массы всех белков 50S субъединицы) такой препарат все еще обладает активностью пептидил-трансферазы. И только целенаправленное повреждение рРНК полностью лишает его этой активности (точно так же, как оно лишает активности целую 50S субъединицу).

Отталкиваясь от этого результата, удалось подобрать условия обратного эксперимента, в котором 23S рРНК, синтезированная *in vitro*, оказалась способной к катализу образования пептидной связи между ацетил-фенилаланил-тРНК и фенилаланил-тРНК. Выход продукта, Ac-Phe-Phe, получается очень низким, говоря о том, что для высокой эффективности катализа 23S рРНК нуждается в поддержке белков. Но, принимая во внимание, что основополагающая каталитическая активность принадлежит именно РНК, роль белков, должно быть, сводится к приданию рРНК правильной пространственной структуры либо к предоставлению субстратов и т. п. Аналогичную реакцию можно осуществить, взяв вместо полноразмерной синтетической 23S рРНК смесь синтетических доменов 23S рРНК. Считается, что каталитический центр 23S рРНК приходится на домен V, хотя сам по себе этот домен (фрагмент РНК) обладает уже совсем небольшой каталитической активностью. Мутации по рибонуклеотидному положению 2252 домена V 23S рРНК «убивают» пептидил-трансферазную активность рибосомы; эта область 23S рРНК входит в состав сайта P.

Кристаллическая структура 50S субъединицы архебактерий подтвердила, что активный центр пептидилтрансферазы состоит в основном из 23S рРНК. Мало того, в радиусе 18 Å от активного центра, где происходит реакция переноса (с образованием пептидной связи), нет вообще никаких белков!

Для образования пептидной связи требуется атака аминогруппы одной аминокислоты на карбоксильную группу другой аминокислоты. Для эффективного катализа нужно, чтобы остаток основания принял на себя атом водорода, отпускаемый аминогруппой (как показано на **РИС. 8.48**). Если катализатором этой реакции действительно служит рРНК, именно она должна предоставлять

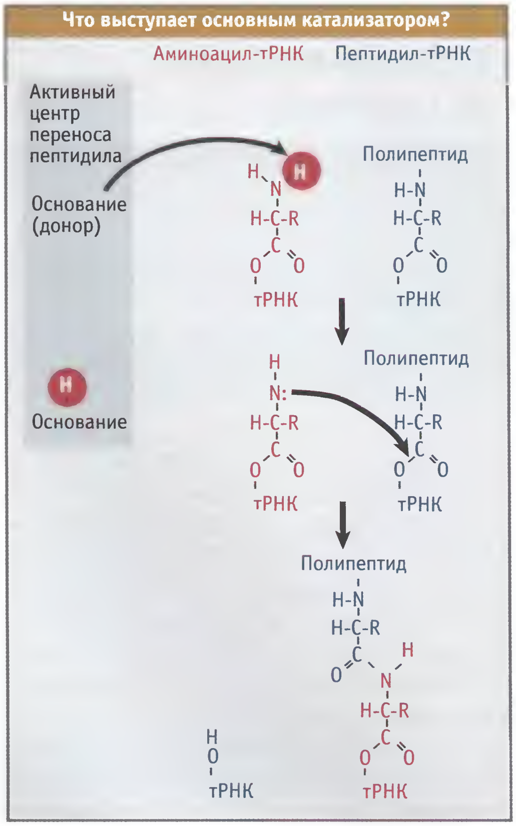


РИС. 8.48 Образование пептидной связи требует кислотно-щелочного катализа с переносом атома водорода на остаток основания

остаток основания, но каким образом это происходит, в точности не известно. Пуриновые и пиримидиновые основания РНК при физиологических значениях pH не являются основаниями в химическом смысле. В принципе, эту роль могло бы играть крайне консервативное азотистое основание (в положении 2451 у *E. coli*), однако его свойства тоже не соответствуют этой роли, да и для пептидилтрансферазной активности его присутствие, кажется, не очень существенно.

Рибосомальные белки, связывающие 23S рРНК за пределами ее каталитически активной области, нужны, скорее всего, для придания ей должной структуры, оптимальной для катализа *in vivo*. Сама идея, что 23S рРНК является каталитическим компонентом рибосомы, полностью согласуется с данными исследований, обсуждаемых в гл. 26 «Сплайсинг и процессинг РНК»; в этих исследованиях были доказаны каталитические свойства РНК, проявляемые в ходе процессинга — превращения продукта транскрипции в субстрат трансляции. (Для генов эукариот процессинг транскриптов выделяется в отдельный этап экспрессии.) Эти данные косвенно подкрепляют представление о том, что функция рибосомы — это эволюционная производная от функции ныне исчезнувшего класса свободных каталитических РНК.

8.20 Встреча большой и малой субъединиц меняет их структуру

Ключевые понятия:

- При образовании полной рибосомы головка 30S субъединицы совершает поворот на шее.
- В полных рибосомах активный центр пептидил-трансферазы в составе 50S субъединицы работает гораздо эффективнее, чем в свободных 50S субъединицах.
- В области взаимодействия 30S и 50S субъединиц есть множество участков, доступных растворителю.

Множество данных косвенно свидетельствует о том, что структура свободных субъединиц рибосомы довольно сильно меняется при их объединении в полную рибосому. Одним из главных индикаторов этого эффекта служит изменение паттерна чувствительности рРНК к внешним агентам (об интерпретации таких паттернов см. в разд. 18.18 «16S рРНК играет активную роль в синтезе белка»). Сравнение кристаллических структур разобщенных субъединиц со структурой полной рибосомы (несмотря на меньшее разрешение последней) позволяет констатировать наличие существенных различий. Анализ кристаллической структуры полной рибосомы *E. coli* с максимальным на сегодняшний день разрешением в 3,5 Å доказывает существование двух различных конформаций рибосомы, вероятно, соответствующих разным этапам синтеза белка.

Кристалл содержит по две рибосомы на ячейку решетки, и у каждой из них своя конформация. Отличия конформаций выражают изменения в расположении структурных массивов внутри каждой субъединицы; самое яркое отличие состоит в том, что в одной из конформаций гловка 30S субъединицы совершает поворот вокруг шеи на 6° в направлении сайта E. Поворот на 6° в противоположном направлении был зафиксирован (при меньшем разрешении) для структуры рибосом *Thermus thermophilus*, связанных с мРНК и содержащих тРНК в сайтах A и P. Все вместе это говорит о том, что «голова» малой субъединицы способна поворачиваться с суммарной амплитудой 12°. Повороты головки как бы провожают каждую тРНК в ее пути по рибосоме. Рибосома хотя бы «мотает головой», контролируя движения мРНК и тРНК.

Конформационные изменения, происходящие при объединении субъединиц в полную рибосому, в малой субъединице выражены гораздо сильнее. Скорее всего, эти изменения важны для контроля положения и движения мРНК. В большой субъединице самое серьезное конформационное изменение затрагивает активный центр пептидил-трансферазы. Свободные 50S субъединицы катализируют образование пептидной связи в 10^3 раз хуже, чем полные рибосомы, — вероятно, субстратам этой реакции гораздо «удобнее» располагаться в активном центре полной рибосомы.

Одна из отличительных особенностей структуры полной рибосомы — очень высокая плотность контактов с растворителем (в который погружена рибосома) в области соприкосновения субъединиц. Доступность растворителя (воды) может помогать образованию и разрыву контактов при ассоциации и диссоциации субъединиц и, кроме того, может облегчать деформацию рибосомы во время транслокации.

8.21 Резюме

Рибосомы — это рибонуклеопротеиновые частицы, бóльшую часть массы которых составляет рРНК. У всех рибосом форма в целом одна и та же (бактериальные 70S рибосомы изучены лучше, чем все остальные). Малая, 30S, субъединица бактериальной рибосомы имеет сплюснутую форму, в ней хорошо различимы тело (на него приходится около двух третей массы субъединицы), и головка, разделенные шеей. Положения белков в составе малой субъединицы известны лишь приблизительно. Большая, 50S, субъединица бактериальной рибосомы более или менее шаровидна; она имеет несколько выступов; самый заметный из них — центральный протуберанец.

Каждая субъединица рибосомы содержит одну большую молекулу рРНК (у прокариот это 16S и 23S; в цитоплазме эукариот — 18S и 28S рРНК). Есть также несколько малых рРНК, самая известная из которых — 5S рРНК большой субъединицы бактериальной рибосомы. Большие рРНК имеют выраженную вторичную структуру, состоящую из множества коротких несовершенных дуплексов, перемежаемых одноцепочечными петлями. Эволюционно консервативные особенности рРНК вычисляют методом сравнения рибонуклеотидных последовательностей и вторичных структур для разных видов организмов. В 16S рРНК насчитывается четыре четко выраженных домена; в 23S рРНК их шесть. В эукариотических рРНК есть дополнительные домены.

Анализ кристаллической структуры рибосомы показывает асимметрию в распределении РНК и белка в составе 30S субъединицы. РНК сосредоточена у поверхности соприкосновения с 50S субъединицей; со стороны 50S субъединицы эта поверхность образована белком, прошитым длинными «балками» двухцепочечной РНК. Объединение 30S субъединицы с 50S субъединицей в полную рибосому происходит во многом за счет контактов между 16S и 23S рРНК. На поверхности раздела субъединиц в полной рибосоме есть множество точек, доступных растворителю. Когда субъединицы встречаются, чтобы образовать полную рибосому, в каждой из них происходят определенные структурные изменения.

В обеих субъединицах рибосомы есть несколько активных сайтов; они сгруппированы в так называемом трансляционном отделе рибосомы, где, собственно, и происходит синтез белка. Растущий полипептид

покидает рибосому через так называемый район выхода, который может быть связан с мембраной. Главными активными сайтами рибосомы являются сайты P и A, сайт E, сайты связывания факторов EF-Tu и EF-G, активный центр пептидил-трансферазы, а также сайт связывания мРНК. Конформация рибосомы очень динамична; она изменяется согласно требованиям каждого этапа синтеза; во всяком случае, доступность отдельных участков больших рРНК в ходе синтеза белка непрерывно изменяется.

Сайты P и A содержат по одной тРНК, которые расположены параллельно друг другу. Их антикодоновые петли взаимодействуют с мРНК в бороздке на поверхности 30S субъединицы. Оставшиеся части обеих тРНК удерживаются 50S субъединицей. Чтобы их аминоацильные концы пришли в сближение, тРНК в сайте A должна поменять свою конформацию. Активный центр пептидил-трансферазы находится на стыке тРНК-связывающих сайтов P и A; он состоит из 23S рРНК, которая обладает собственной каталитической активностью пептидил-трансферазы (хотя для поддержания каталитически активной структуры ей, по-видимому, необходима помощь целого ряда рибосомальных белков).

Об активной роли рРНК в синтезе белка говорит тот факт, что многие точечные мутации в рРНК парализуют функцию рибосомы. Другим подтверждением ее активной роли служат взаимодействия рРНК с мРНК и тРНК, которые можно детектировать при помощи химической фиксации. рРНК служит мишенью некоторых антибиотиков и других агентов, ингибирующих синтез белка. Наконец, об активной роли рРНК в синтезе белка можно судить по образованию гибридных двухцепочечных участков между рРНК и мРНК, а также между рРНК и тРНК, и принципиальной функциональной важности этих преходящих гибридных дуплексов. Так, во время инициации трансляции 3′-концевая область 16S рРНК взаимодействует с комплементарным сегментом мРНК — последовательностью Шайна–Дальгарно в сайте инициации; внутренние области рРНК образуют ряд контактов с тРНК в сайтах P и A.

Любой значащий кодон мРНК узнается молекулой аминоацил-тРНК, содержащей антикодон, комплементарный данному кодону, и несущей аминокислоту, соответствующую этому кодону. Особая инициаторная тРНК (fMet-тРНК$_f$ у прокариот, Met-тРНК$_i$ в цитоплазме эукариот) узнает инициаторный кодон AUG, с которого начинается почти любая кодирующая последовательность (у прокариот иногда используются инициаторные кодоны GUG и UUG). Еще три кодона не узнаются никакими аминоацил-тРНК; это терминирующие, или нонсенс-кодоны, UAA, UAG и UGA.

Рибосомы, завершившие синтез белка, пополняют собой клеточный пул свободных рибосом, который уравновешен пулом разобщенных больших и малых субъединиц. Малые субъединицы связывают мРНК; после этого к ним присоединяются большие субъединицы; образовавшиеся полные рибосомы синтезируют новые полипептиды. Узнавание прокаритического сайта инициации (сайта связывания рибосомы) в мРНК — сложный процесс с участием опять же рРНК: 3′-концевая область 16S рРНК в малой субъединице связывает последовательность Шайна–Дальгарно, предваряющую собой инициаторный кодон. Узнавание эукариотического сайта инициации происходит иначе: сначала малая субъединица узнает мРНК как молекулу данного класса, по кэпу на 5′-конце (модификация, присущая всем эукариотическим мРНК). Связав кэп, малая субъединица начинает двигаться вдоль мРНК в поисках инициаторного кодона AUG. Когда она встречает подходящий AUG (чаще всего это первый AUG на ее пути), к ней присоединяется большая субъединица, и полная рибосома приступает к синтезу полипептида.

В рибосоме одновременно помещаются две тРНК. В сайте P находится полипептидил-тРНК; перед тем, как покинуть сайт P, она передает свой полипептидил на тРНК, расположенную в сайте A; этот сайт служит рибосоме для конкурентного связывания тРНК, несущих новые аминокислоты растущего полипептида. В бактериальных рибосомах есть дополнительный сайт E, через который, покидая рибосому, проходят использованные деацилированные тРНК. В каждом рабочем цикле элонгации образуется новая пептидная связь: полипептидная цепочка переносится из сайта P на аминоацил в сайте A; при этом она удлиняется на один аминокислотный остаток. тРНК в сайте A превращается из аминоацил-тРНК в полипептидил-тРНК, тогда как в сайте P остается покинутая растущим полипептидом, деацилированная тРНК.

В каждом рабочем цикле элонгации за актом образования пептидной связи следует акт транслокации. Во время транслокации деацилированная тРНК переходит в сайт E, а пептидил-тРНК переходит из сайта A в сайт P. Катализатором транслокации выступает фактор элонгации EF-G; для этой реакции (равно как и для обеспечения ряда других моментов работы рибосомы) необходим гидролиз ГТФ. При транслокации рибосома, деформируясь, переходит в некое промежуточное (гибридное) состояние, с характерным сдвигом субъединиц друг относительно друга.

Синтез белка обходится недешево. Различные этапы синтеза обеспечиваются энергией за счет гидролиза АТФ либо ГТФ. Так, для «зарядки» тРНК аминокислотами или для расплетания вторичной структуры мРНК необходим гидролиз АТФ. По некоторым оценкам, на синтез полипептидов расходуется до 90% всех молекул АТФ, которыми располагают быстро растущие бактерии.

На каждом этапе синтеза белка рибосома нуждается в помощи дополнительных факторов, совершающих циклы ассоциации и диссоциации с рибосомой. Факторы IF участвуют в инициации синтеза белка (инициации трансляции мРНК) у прокариот. Фактор IF-3 нужен для связывания 30S субъединиц с мРНК и поддержания 30S субъединиц в свободном состоянии, необходимом для инициации. Фактор IF-2 катализиру-

ет связывание fMet-тРНК$_f$ с 30S субъединицей и не допускает другие тРНК в инициацию. После связывания инициаторной тРНК с комплексом инициации происходит гидролиз ГТФ. Чтобы большая субъединица рибосомы могла примкнуть к комплексу инициации, факторы инициации должны покинуть комплекс.

В эукариотической инициации участвует гораздо большее число факторов. Некоторые из них содействуют первоначальному связыванию 40S субъединицы с кэпированным 5′-концом мРНК, другие факторы в это время приносят инициаторную тРНК. После осуществления первоначального связывания с 5′-концом малая субъединица сканирует мРНК до тех пор, пока не встретит подходящий кодон AUG. На этом факторы инициации получают свободу; вместо них в комплекс инициации вступает 60S субъединица.

У прокариот нормальный ход элонгации обеспечивается факторами EF. Фактор EF-Tu помещает аминоацил-тРНК в 70S рибосому; если он справился с этой задачей, ему разрешается гидролизовать свой ГТФ и покинуть рибосому для восстановления активной формы и «поимки» новой тРНК. Восстановить активную форму EF-Tu помогает фактор EF-Ts. Еще один фактор, EF-G, участвует в транслокации. Присутствие в рибосоме EF-Tu исключает присутствие EF-G, и наоборот; такое взаимное исключение двух факторов служит гарантией строгого чередования актов образования пептидной связи с актами транслокации.

Терминация синтеза происходит по одному из трех специальных кодонов (UAA, UAG и UGA). Факторы RF класса 1, узнавая эти стоп-кодоны, побуждают рибосому к гидролизу пептидил-тРНК. Чтобы после этого RF класса 1 мог покинуть рибосому, нужен дополнительный фактор, RF класса 2 (RF3 у *E. coli*). Все ГТФ-зависимые факторы синтеза белка (ГТФ-зависимые факторы трансляции IF-2, EF-Tu, EF-G, RF3) похожи друг на друга. Факторы EF-G и RF3 как бы подражают РНК-белковой структуре IF-2 или EF-Tu в комплексе с тРНК. По последним данным, все ГТФ-зависимые факторы трансляции вписываются в один и тот же активный сайт рибосомы — так называемый сайт связывания G-факторов.

Литература

8.4 В инициации синтеза белка у бактерий участвуют 30S субъединицы и вспомогательные факторы

Обзор

Maitra, U., (1982). Initiation factors in protein biosynthesis. *Annu. Rev. Biochem.* 51, 869–900.

Исследование

Carter, A. P., Clemons, W. M., Brodersen, D. E., Morgan-Warren, R. J., Hartsch, T., Wimberly, B. T., and Ramakrishnan, V. (2001). Crystal structure of an initiation factor bound to the 30S ribosomal subunit. *Science* 291, 498–501.

Dallas, A. and Noller, H. F. (2001). Interaction of translation initiation factor 3 with the 30S ribosomal subunit. *Mol. Cell* 8, 855–864.

Moazed, D., Samaha, R. R., Gualerzi, C., and Noller, H. F. (1995). Specific protection of 16S rRNA by translational initiation factors. *J. Mol. Biol.* 248, 207–210.

8.5 Особая инициаторная тРНК закладывает первое звено полипептида

Исследование

Lee, C. P., Seong, B. L., and RajBhandary, U. L. (1991). Structural and sequence elements important for recognition of *E. coli* formylmethionine tRNA by methionyl-tRNA transformylase are clustered in the acceptor stem. *J. Biol. Chem.* 266, 18012–18017.

Marcker, K. and Sanger, F. (1964). N-Formyl-methionyl-S-RNA. *J. Mol. Biol.* 8, 835–840.

Sundari, R. M., Stringer, E. A., Schulman, L. H., and Maitra, U. (1976). Interaction of bacterial initiation factor 2 with initiator tRNA. *J. Biol. Chem.* 251, 3338–3345.

8.8 У эукариот малые субъединицы рибосом сканируют мРНК в поисках сайтов инициации

Обзор

Hellen, C. U. and Sarnow, P. (2001). Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev.* 15, 1593–1612.

Kozak, M. (1978). How do eukaryotic ribosomes select initiation regions in mRNA? *Cell* 15, 1109–1123.

Kozak, M. (1983). Comparison of initiation of protein synthesis in prokaryotes, eukaryotes, and organelles. *Microbiol. Rev.* 47, 1–45.

Исследование

Kaminski, A., Howell, M. T., and Jackson, R. J. (1990). Initiation of encephalomyocarditis virus RNA translation: the authentic initiation site is not selected by a scanning mechanism. *EMBO J.* 9, 3753–3759.

Pelletier, J. and Sonenberg, N. (1988). Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* 334, 320–325.

Pestova, T. V., Hellen, C. U., and Shatsky, I. N. (1996). Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. *Mol. Cell Biol.* 16, 6859–6869.

Pestova, T. V., Shatsky, I. N., Fletcher, S. P., Jackson, R. J., and Hellen, C. U. (1998). A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. *Genes Dev.* 12, 67–83.

8.9 У эукариот множество факторов инициации объединено в комплекс

Обзор

Dever, T. E. (2002). Gene-specific regulation by general translation factors. *Cell* 108, 545–556.

Gingras, A. C., Raught, B., and Sonenberg, N. (1999). eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu. Rev. Biochem.* 68, 913–963.

Hershey, J. W. B. (1991). Translational control in mammalian cells. *Annu. Rev. Biochem.* 60, 717–755.

Merrick, W. C. (1992). Mechanism and regulation of eukaryotic protein synthesis. *Microbiol. Rev.* 56, 291–315.

Pestova, T. V., Kolupaeva, V. G., Lomakin, I. B., Pilipenko, E. V., Shatsky, I. N., Agol, V. I., and Hellen, C. U. (2001). Molecular mechanisms of translation initiation in eukaryotes. *Proc. Natl. Acad. Sci. USA* 98, 7029–7036.

Sachs, A., Sarnow, P., and Hentze, M. W. (1997). Starting at the beginning, middle, and end: translation initiation in eukaryotes. *Cell* 89, 831–838.

Исследование

Asano, K., Clayton, J., Shalev, A., and Hinnebusch, A. G. (2000). A multifactor complex of eukaryotic initiation factors, eIF1, eIF2, eIF3, eIF5, and initiator tRNA(Met) is an important translation initiation intermediate *in vitro*. *Genes Dev.* 14, 2534–2546.

Huang, H. K., Yoon, H., Hannig, E. M., and Donahue, T. F. (1997). GTP hydrolysis controls stringent selection of the AUG start codon during translation initiation in *S. cerevisiae*. *Genes Dev.* 11, 2396–2413.

Pestova, T. V., Lomakin, I. B., Lee, J. H., Choi, S. K., Dever, T. E., and Hellen, C. U. (2000). The joining of ribosomal subunits in eukaryotes requires eIF5B. *Nature* 403, 332–335.

Tarun, S. Z. and Sachs, A. B. (1996). Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *EMBO J.* 15, 7168–7177.

8.12 При транслокации рибосома приходит в движение

Обзор

Ramakrishnan, V. (2002). Ribosome structure and the mechanism of translation. *Cell* 108, 557–572.

Wilson, K. S. and Noller, H. F. (1998). Molecular movement inside the translational engine. *Cell* 92, 337–349.

Исследование

Moazed, D., and Noller, H. F. (1986). Transfer RNA shields specific nucleotides in 16S ribosomal RNA from attack by chemical probes. *Cell* 47, 985–994.

Moazed, D. and Noller, H. F. (1989). Intermediate states in the movement of tRNA in the ribosome. *Nature* 342, 142–148.

8.13 Факторы элонгации поочередно связываются с рибосомой

Исследование

Nissen, P., Kjeldgaard, M., Thirup, S., Polekhina, G., Reshetnikova, L., Clark, B. F., and Nyborg, J. (1995). Crystal structure of the ternary complex of Phe-tRNAPhe, EF-Tu, and a GTP analog. *Science* 270, 1464–1472.

Stark, H., Rodnina, M. V., Wieden, H. J., van Heel, M., and Wintermeyer, W. (2000). Large-scale movement of elongation factor G and extensive conformational change of the ribosome during translocation. *Cell* 100, 301–309.

8.15 Терминирующие кодоны узнаются белковыми факторами

Обзор

Eggertsson, G. and Soll, D. (1988). Transfer RNA-mediated suppression of termination codons in *E. coli*. *Microbiol. Rev.* 52, 354–374.

Frolova, L. et al. (1994). A highly conserved eukaryotic protein family possessing properties of polypeptide chain release factor. *Nature* 372, 701–703.

Nissen, P., Kjeldgaard, M., and Nyborg, J. (2000). Macromolecular mimicry. *EMBO J.* 19, 489–495.

Исследование

Freistroffer, D. V., Kwiatkowski, M., Buckingham, R. H., and Ehrenberg, M. (2000). The accuracy of codon recognition by polypeptide release factors. *Proc. Natl. Acad. Sci. USA* 97, 2046–2051.

Ito, K., Ebihara, K., Uno, M., and Nakamura, Y. (1996). Conserved motifs in prokaryotic and eukaryotic polypeptide release factors: tRNA-protein mimicry hypothesis. *Proc. Natl. Acad. Sci. USA* 93, 5443–5448.

Klaholz, B. P., Myasnikov, A. G., and Van Heel, M. (2004). Visualization of release factor 3 on the ribosome during termination of protein synthesis. *Nature* 427, 862–865.

Mikuni, O., Ito, K., Moffat, J., Matsumura, K., McCaughan, K., Nobukuni, T., Tate, W., and Nakamura, Y. (1994). Identification of the *prfC* gene, which encodes peptide-chain-release factor 3 of *E. coli*. *Proc. Natl. Acad. Sci. USA* 91, 5798–5802.

Milman, G., Goldstein, J., Scolnick, E., and Caskey, T. (1969). Peptide chain termination. 3. Stimulation of *in vitro* termination. *Proc. Natl. Acad. Sci. USA* 63, 183–190.

Scolnick, E. et al. (1968). Release factors differing in specificity for terminator codons. Proc. Natl. Acad. Sci. USA 61, 768–774.

Selmer, M., Al-Karadaghi, S., Hirokawa, G., Kaji, A., and Liljas, A. (1999). Crystal structure of *Thermotoga maritima* ribosome recycling factor: a tRNA mimic. *Science* 286, 2349–2352.

Song, H., Mugnier, P., Das, A. K., Webb, H. M., Evans, D. R., Tuite, M. F., Hemmings, B. A., and Barford, D. (2000). The crystal structure of human eukaryotic release factor eRF1—mechanism of stop codon recognition and peptidyl-tRNA hydrolysis. *Cell* 100, 311–321.

8.16 Обе субъединицы рибосомы пронизаны рибосомальными РНК

Обзор

Hill, W. E. et al. (1990). *The Ribosome.* Washington, DC: American Society for Microbiology.

Noller, H. F. (1984). Structure of ribosomal RNA. *Annu. Rev. Biochem.* 53, 119–162.

Noller, H. F. (2005). RNA structure: reading the ribosome. *Science* 309, 1508–1514.

Noller, H. F. and Nomura, M. (1987). *E. coli* and *S. typhimurium*. Washington, DC: American Society for Microbiology.

Wittman, H. G. (1983). Architecture of prokaryotic ribosomes. *Annu. Rev. Biochem.* 52, 35–65.

Исследование

Ban, N., Nissen, P., Hansen, J., Capel, M., Moore, P. B., and Steitz, T. A. (1999). Placement of protein and RNA structures into a 5 Å-resolution map of the 50S ribosomal subunit. *Nature* 400, 841–847.

Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289, 905–920.

Clemons, W. M. et al. (1999). Structure of a bacterial 30S ribosomal subunit at 5.5 Å resolution. *Nature* 400, 833–840.

Wimberly, B. T., Brodersen, D. E., Clemons W. M. Jr., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000). Structure of the 30S ribosomal subunit. *Nature* 407, 327–339.

Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, A., Earnest, T. N., Cate, J. H. D., and Noller, H. F. (2001). Crystal structure of the ribosome at 5.5 Å resolution. *Science* 292, 883–896.

8.17 В рибосоме есть несколько активных центров

Обзор

Lafontaine, D. L. and Tollervey, D. (2001). The function and synthesis of ribosomes. *Nat. Rev. Mol. Cell Biol.* 2, 514–520.

Moore, P. B. and Steitz, T. A. (2003). The structural basis of large ribosomal subunit function. *Annu. Rev. Biochem.* 72, 813–850.

Ramakrishnan, V. (2002). Ribosome structure and the mechanism of translation. *Cell* 108, 557–572.

Исследование

Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N., and Noller, H. F. (1999). X-ray crystal structures of 70S ribosome functional complexes. *Science* 285, 2095–2104.

Fredrick, K., and Noller, H. F. (2003). Catalysis of ribosomal translocation by sparsomycin. *Science* 300, 1159–1162.

Sengupta, J., Agrawal, R. K., and Frank, J. (2001). Visualization of protein S1 within the 30S ribosomal subunit and its interaction with messenger RNA. *Proc. Natl. Acad. Sci. USA* 98, 11991–11996.

Simonson, A. B. and Simonson, J. A. (2002). The transorientation hypothesis for codon recognition during protein synthesis. *Nature* 416, 281–285.

Valle, M., Sengupta, J., Swami, N. K., Grassucci, R. A., Burkhardt, N., Nierhaus, K. H., Agrawal, R. K., and Frank, J. (2002). Cryo-EM reveals an active role for aminoacyl-tRNA in the accommodation process. *EMBO J.* 21, 3557–3567.

Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, A., Earnest, T. N., Cate, J. H. D., and Noller, H. F. (2001). Crystal structure of the ribosome at 5.5 Å resolution. *Science* 292, 883–896.

8.18 16S рРНК играет активную роль в синтезе белка

Обзор

Noller, H. F. (1991). Ribosomal RNA and translation. *Annu. Rev. Biochem.* 60, 191–227.

Yonath, A. (2005). Antibiotics targeting ribosomes: resistance, selectivity, synergism and cellular regulation. *Annu. Rev. Biochem.* 74, 649–679.

Исследование

Lodmell, J. S. and Dahlberg, A. E. (1997). A conformational switch in *E. coli* 16S rRNA during decoding of mRNA. *Science* 277, 1262–1267.

Moazed, D., and Noller, H. F. (1986). Transfer RNA shields specific nucleotides in 16S ribosomal RNA from attack by chemical probes. *Cell* 47, 985–994.

Yoshizawa, S., Fourmy, D., and Puglisi, J. D. (1999). Recognition of the codon-anticodon helix by rRNA. *Science* 285, 1722–1725.

8.19 23S рРНК обладает пептидилтрансферазной активностью

Исследование

Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289, 905–920.

Bayfield, M. A., Dahlberg, A. E., Schulmeister, U., Dorner, S., and Barta, A. (2001). A conformational change in the ribosomal peptidyl transferase center upon active/inactive transition. *Proc. Natl. Acad. Sci. USA* 98, 10096–10101.

Noller, H. F., Hoffarth, V., and Zimniak, L. (1992). Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* 256, 1416–1419.

Samaha, R. R., Green, R., and Noller, H. F. (1995). A base pair between tRNA and 23S rRNA in the peptidyl transferase center of the ribosome. *Nature* 377, 309–314.

Thompson, J., Thompson, D. F., O'Connor, M., Lieberman, K. R., Bayfield, M. A., Gregory, S. T., Green, R., Noller, H. F., and Dahlberg, A. E. (2001). Analysis of mutations at residues A2451 and G2447 of 23S rRNA in the peptidyltransferase active site of the 50S ribosomal subunit. *Proc. Natl. Acad. Sci. USA* 98, 9002–9007.

8.20 Встреча большой и малой субъединиц меняет их структуру

Ссылка

Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., and Cate, J. H. (2005). Structures of the bacterial ribosome at 3.5 Å resolution. *Science* 310, 827–834.

Расшифровка генетического кода

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

9.1 Введение

9.2 Схожие кодоны кодируют аминокислоты с похожими свойствами

- 61 из 64 возможных кодонов кодируют 20 аминокислот.
- 3 кодона не кодируют аминокислоты, а являются стоп-кодонами, вызывая терминацию синтеза белка.
- Генетический код универсален и работает одинаково в организмах разного уровня сложности.
- Одной и той же аминокислоте может соответствовать несколько кодонов.
- Кодоны, кодирующие одну и ту же аминокислоту, похожи.
- Тенденция к сходству кодонов для похожих по свойствам аминокислот существенно снижает эффект мутаций.

9.3 Кодон-антикодоновое распознавание допускает неоднозначное спаривание

- Триплеты, кодирующие одну и ту же аминокислоту, часто различаются только нуклеотидным основанием в третьей позиции.
- Необязательность соответствия оснований в первом положении антикодона и третьем положении кодона является результатом особой структуры антикодоновой петли.

9.4 тРНК образуется из длинного предшественника

- Зрелая тРНК образуется в результате процессинга предшественника.
- 5′-конец зрелой тРНК образуется в результате расщепления молекулы-предшественника эндонуклеазой — РНКазой P.
- 3′-конец образуется в результате расщепления, которое следует за удалением нескольких последних оснований и добавлением общего для всех тРНК терминального триплета CCA.

9.5 тРНК содержит модифицированные основания

- тРНК содержит более 50 модифицированных оснований.
- Модификация обычно представляет собой изменение исходных нуклеотидов тРНК однако в некоторых случаях, заключается в замене одних оснований на другие.

9.6 Модифицированные основания влияют на кодон-антикодоновые взаимодействия

- Модификации нуклеотидов антикодона влияют на набор воблирующих оснований, тем самым определяя специфичность тРНК.

9.7 Изменения в универсальном генетическом коде носят спорадический характер

- У некоторых видов встречаются отклонения от универсального генетического кода.
- Изменения в генетическом коде наиболее характерны для генома митохондрий, на основании этих изменений можно построить филогенетическое древо.
- В ядерном геноме отклонения от стандарта крайне редки и обычно затрагивают лишь стоп-кодоны.

9.8 Некоторые стоп-кодоны могут встраивать в белок стандартные аминокислоты

- Ряд генов санкционирует особое прочтение специфичных кодонов.
- Взаимодействие селено-Cys-тРНК и UGA-кодона требует участия нескольких белков, модифицирующих Cys-тРНК и присоединяющих ее к рибосоме.
- Пирролизин может транслироваться с UAG-кодона.

9.9 Синтетазы присоединяют аминокислоты к тРНК

- Аминоацил-тРНК-синтетазы — ферменты, присоединяющие аминокислоты к молекулам тРНК. Образование аминоацил-тРНК идет в 2 стадии с использованием энергии АТФ.
- В каждой клетке присутствует 20 аминоацил-тРНК-синтетаз. Каждой аминокислоте соответствует как минимум одна синтетаза, которая обнаруживает все тРНК, представляющие эту аминокислоту.
- Узнавание молекулой тРНК своей аминокислоты обеспечивается небольшим числом сайтов связывания в ее нуклеотидной последовательности.

9.10 Аминоацил-тРНК-синтетазы делят на 2 класса

- В зависимости от аминокислотной последовательности и структурного сходства аминоацил-тРНК-синтетазы делят на 2 класса.

9.11 Синтетазы — ферменты с корректирующей активностью

- Специфичность реакции узнавания между аминокислотой и тРНК контролируется аминоацил-тРНК-синтетазой, корректирующая активность которой обеспечивает исправление ошибок в случае связывания неправильного компонента.

9.12 Мутантные антикодоны дают возможность супрессорным тРНК считывать новые кодоны

- Супрессорые тРНК обычно несут мутации в антикодоне, приводящие к изменению типа кодона, узнаваемого данной тРНК.
- В результате супрессорной мутации в антикодоне тРНК приобретает способность распознавать стоп-кодон как смысловой, обеспечивая включение аминокислоты в полипептидную цепь и синтез белка за стоп-кодон. Это явление получило название нонсенс-супрессии или супрессии преждевременной терминации белкового синтеза.
- Миссенс-супрессия проявляется в способности тРНК узнавать кодоны, отличные от обычных для данной тРНК и приводит к замене одной аминокислоты на другую.

9.13 Для каждого терминирующего кодона существуют свои нонсенс-супрессоры

- Каждому типу нонсенс-кодона соответствует тРНК с определенным мутантным антикодоном.
- Иногда супрессорные мутации возникают не только в антикодонах, но и в других участках молекулы тРНК.

9.14 Супрессорные тРНК могут конкурировать с тРНК дикого типа за связывание с кодонами

- Супрессорные тРНК и тРНК дикого типа, обладающие одинаковыми антикодонами, конкурируют друг с другом за возможность взаимодействия с соответствующим кодоном.
- Негативный эффект супрессии заключается в считывании и правильных стоп-кодонов.
- Кодон UGA даже в отсутствие мутантной супрессорной тРНК может узнаваться Trp-тРНК с частотой 1–3%.

9.15 Рибосомы влияют на точность трансляции

- Структура 16s-рРНК в рибосомальных сайтах А и Р влияет на точность трансляции.

9.16 Перекодирование меняет значение кодона

- Изменение смысла кодона может быть следствием возникновения мутантных тРНК или тРНК со специфичными свойствами.
- В результате сдвига или пропуска фрагмента рибосома может менять рамку считывания, оба этих процесса зависят от структуры и свойств мРНК.

9.17 Сдвиг рамок считывания происходит на «скользких» последовательностях

- Рамка считывания определяется последовательностью нуклеотидов в мРНК и рибосомальным окружением.
- «Скользкие» последовательности способствуют сдвигу тРНК, спаренной с антикодоном на одно основание, изменяя тем самым рамку считывания.
- Трансляция некоторых мРНК зависит от регулярных «запрограммированных» сдвигов рамки считывания.

9.18 «Перепрыгивание» через основания осуществляется перемещением рибосомы

9.19 Резюме

9.1 Введение

Последовательность кодирующей цепи ДНК (в направлении от 5′- к 3′-концу) представлена нуклеотидными триплетами (кодонами), соответствующими определенным аминокислотам в последовательности белка (синтезируемого соответственно от N-конца к C-концу). Секвенирование ДНК и белков делает возможным прямое сопоставление последовательности нуклеотидов и аминокислот. Из 64 возможных кодонов (каждый из четырех нуклеотидов может занимать одну из трех позиций в кодоне, формируя $4^3 = 64$ различных комбинаций) 61 триплет кодирует аминокислоту (смысловые кодоны), а три (стоп-кодоны) вызывают терминацию синтеза белка.

Соответствие смыслового кодона определенной аминокислоте детерминировано геном тРНК, несущей ее. Однако для стоп-кодонов нет соответствующих тРНК, а терминация трансляции опосредована белковыми факторами.

Первые исследования структуры генетического кода показали, что генетическая информация хранится в форме нуклеотидных триплетов, однако соответствие кодонов аминокислотам оставалось невыясненным. С целью определения даного соответствия еще до появления секвенирования было разработано два метода, с помощью которых и расшифровали большинство кодонов. Систему трансляции синтетических полинуклеотидов стали применять для синтеза белка после того, как в 1961 г. Ниренберг показал, что полиуридиловая кислота (poly(U)), как и клеточные мРНК, может служить матрицей для синтеза белка. При этом из фенилаланина синтезируется полифениланин. Из этого следовало, что кодон UUU соответствует фенилаланину. Второй метод, разработанный позже, заключался в том, что тринуклеотиды (которые уже умели синтезировать с заданной последовательностью) стимулируют связывание специфичных к данному триплету молекул аминоацил-тРНК с рибосомами (имитируя процесс взаимодействия кодона с антикодоном на рибосоме). Определяя, какая именно аминокислота входит в состав связавшейся с рибосомой аминоацил-тРНК, биохимики установили значение каждого из триплетов. Два описанных выше метода позволили точно определить значение всех смысловых кодонов.

Как уже было сказано, 61 из 64 возможных кодонов кодируют аминокислоты, а 3 оставшихся являются стоп-кодонами, терминирующими синтез белка. Важно, что основание в третьем положении многих кодонов несущественно для их «спаривания» с аминоацил-тРНК. Более того, аминокислоты со схожими биохимическими свойствами кодируются схожими кодонами.

9.2 Схожие кодоны кодируют аминокислоты с похожими свойствами

Ключевые понятия:

- 61 из 64 возможных кодонов кодируют 20 аминокислот.
- 3 кодона не кодируют аминокислоты, а являются стоп-кодонами, вызывая терминацию синтеза белка.
- Генетический код универсален и работает одинаково в организмах разного уровня сложности.
- Одной и той же аминокислоте может соответствовать несколько кодонов.
- Кодоны, кодирующие одну и ту же аминокислоту, похожи.
- Тенденция к сходству кодонов для похожих по свойствам аминокислот существенно снижает эффект мутаций.

Результат расшифровки генетического кода представлен на **РИС. 9.1**. Поскольку число возможных триплетных комбинаций превышает число аминокислот, почти каждая аминокислота соответствует более чем одному кодону. Исключением являются две аминокислоты — метионин и триптофан, имеющие лишь по одному кодону. Кодирующие одну и ту же аминокислоту кодоны называют **синонимичными**. Поскольку генетический код считывается рибосомой с мРНК, его обычно записывают при помощи оснований РНК: U, C, A и G.

Триплеты, кодирующие одну и ту же аминокислоту (либо аминокислоты со схожими свойствами), имеют схожие нуклеотидные последовательности. Часто основание в третьем положении кодона оказывается несущественным для его специфичности. Одна аминокислота может быть представлена четырьмя кодонами, различающимися по основанию в третьем положении. Иногда различие состоит в предпочтении пурина пиримидину в третьем положении кодона. Меньшая специфичность третьего положения в кодоне называется **вырожденностью третьего основания.** Дешифровка нуклеотидного триплета мРНК в соответствующую аминокислоту белка происходит в результате взаимодействия кодона мРНК и антикодона соответствующей аминоацил-тРНК. Этот процесс протекает в рибосомах: небольшое число комплементарных нуклеотидов триплета само по себе не может обеспечить стабильного кодон-антикодонового взаимодействия, однако в А-сайте рибосомы создаются условия, необходимые для стабилизации комплекса кодон-антикодон. Варианты взаимодействия между кодоном и антикодоном включают в себя не только A-U и G-C взаимодействия. Образование нуклеотидных пар в первых двух положениях кодона всегда происходит по каноническим правилам, однако в третьем положении возможно нестрогое соответствие. В итоге одна аминоацил-тРНК может узнавать более одного кодона. Кроме того, на кодон-антикодоновые взаимодействия могут влиять модифицированные нуклеотидные основания, присутствующие в тРНК непосредственно в антикодоне либо рядом с ним.

Тенденция к сходству кодонов для аминокислот одного типа снижает эффект мутаций. При описанной организации кода случайно возникшая замена основания в ДНК с большей вероятностью приведет к замене на схожую по свойствам аминокислоту в белке (или же замены не произойдет вовсе). К примеру, замена триплета CUC на триплет CUG не оказывает никакого влияния на последовательность аминокислот белка, так как оба кодона соответствуют аминокислоте лейцину. А замена CUU на AUU приводит к замещению лейцина очень похожим на него изолейцином.

На **РИС. 9.2** показана зависимость между числом кодонов для каждой аминокислоты и частотой встречаемости данной аминокислоты в белках (*E. coli*). Видно, что между этими двумя параметрами наблюдается достаточно слабая корреляция. Возникает ощущение, что организация генетического кода не оптимизирована.

Три кодона (UAA, UAG и UGA) не кодируют аминокислоты, а вызывают остановку синтеза белка — терминацию. Один из этих кодонов (UAA) является наиболее частым стоп-кодоном при синтезе белков.

Универсален ли генетический код для всех организмов?

При сравнении последовательностей ДНК с соответствующими белковыми последовательностями оказалось, что во всех бактериальных и эукариотических геномах используются одни и те же значения кодонов. В результате мРНК одного вида может быть правильно прочитана *in vivo* или *in vitro* белок-синтезирующим аппаратом другого вида, а кодоны, используемые в мРНК

Генетический код триплетен

Первое основание	Второе основание: U	C	A	G
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys
U	UUC Phe	UCC Ser	UAC Tyr	UGC Cys
U	UUA Leu	UCA Ser	UAA СТОП	UGA СТОП
U	UUG Leu	UCG Ser	UAG СТОП	UGG Trp
C	CUU Leu	CCU Pro	CAU His	CGU Arg
C	CUC Leu	CCC Pro	CAC His	CGC Arg
C	CUA Leu	CCA Pro	CAA Gln	CGA Arg
C	CUG Leu	CCG Pro	CAG Gln	CGG Arg
A	AUU Ile	ACU Thr	AAU Asn	AGU Ser
A	AUC Ile	ACC Thr	AAC Asn	AGC Ser
A	AUA Ile	ACA Thr	AAA Lys	AGA Arg
A	AUG Met	ACG Thr	AAG Lys	AGG Arg
G	GUU Val	GCU Ala	GAU Asp	GGU Gly
G	GUC Val	GCC Ala	GAC Asp	GGC Gly
G	GUA Val	GCA Ala	GAA Glu	GGA Gly
G	GUG Val	GCG Ala	GAG Glu	GGG Gly

РИС. 9.1 Все триплетные кодоны осмысленны: шестьдесят один кодон соответствуют аминокислотам, три кодона вызывают терминацию синтеза белка (STOP)



РИС. 9.2 **Число кодонов для каждой аминокислоты не обнаруживает тесной связи с частотой использования данной аминокислоты в белках**

одного вида, имеют единый смысл для рибосом и тРНК других видов.

Универсальность кода свидетельствует о том, что в процессе эволюции он сложился очень рано. Возможно, сначала код существовал в примитивном виде, при этом малое число кодонов обозначало сравнительно небольшое число аминокислот; один кодон мог соответствовать любому члену одной группы аминокислот. Более точное значение кодонов и большее число аминокислот могли быть «введены» позже. Сначала для узнавания использовались только первые два из трех оснований в кодоне, роль третьего положения определилась позднее. (Согласно некоторым моделям, первоначально могли существовать некоторые стереохимические взаимодействия между аминокислотами и их кодонами. Затем, по мере развития белок-синтезирующей системы, появились более комплексные системы.)

В некоторый момент времени код был «заморожен» в его современном виде, поскольку любые изменения в уже сложившейся системе кодирования приводили к критическим изменениям в белке, обусловленным аминокислотной заменой. Универсальность кода означает, что он возник еще у примитивных клеток, из которых произошли все ныне живущие организмы.

Исключения из универсальности генетического кода очень редки. Изменения чаще всего касаются стоп-кодонов. К примеру, у *Mycoplasma* кодон UGA кодирует аминокислоту триптофан; у некоторых видов ресничных *Tetrahymena* и *Paramecium* кодоны UAA и UAG используются для кодирования глутамина. Систематические изменения кода характерны только для митохондриальной ДНК (см. разд. 9.7 «Изменения в универсальном генетическом коде носят спорадический характер»).

9.3 Кодон-антикодоновое распознавание допускает неоднозначное спаривание

Ключевые понятия:

- Триплеты, кодирующие одну и ту же аминокислоту, часто различаются только нуклеотидным основанием в третьей позиции.
- Необязательность соответствия оснований в первом положении антикодона и третьем положении кодона является результатом особой структуры антикодоновой петли.

Функция тРНК в механизме синтеза белка заключается в том, чтобы после ассоциации с сайтом А рибосомы узнать нужный кодон на мРНК. Взаимодействие между кодоном мРНК и антикодоном тРНК происходит путем комплементарного спаривания оснований, но по правилам, которые значительно отличаются от классических, описывающих взаимодействие пар G-C и A-U при синтезе мРНК.

Правила взаимодействия между кодоном и соответствующим ему антикодоном можно установить при помощи теста на специфичность связывания тРНК с заранее известным триплетом или путем использования тРНК в системе белкового синтеза *in vitro*.

Структура генетического кода позволяет сделать ряд важных заключений о механизме узнавания кодона. На РИС. 9.3 приведены примеры вырожденности третьего основания: в большинстве случаев оно вовсе не релевантно, либо значение играет лишь принадлежность основания к пуринам или пиримидинам.

По принципу кодирования одной аминокислоты можно выделить восемь семейств кодонов. Все четыре кодона в каждом из семейств совпадают по первым двум основаниям и кодируют одну и ту же аминокислоту. Следовательно, для таких семейств третье основание как бы не играет роли в определении смысла кодона. Аналогично описанным семействам существуют семь пар кодонов, смысл которых не зависит от того, какой из пиримидинов присутствует в третьем положении. Кроме того, имеется еще пять пар кодонов, в которых любой из пуринов может находиться в третьем положении, не изменяя при этом смысл кодона.

Известно только три случая, когда основание, находясь в третьем положении, придает уникальный смысл кодону. AUG (кодирует метионин), UGG (кодирует триптофан) и UGA (стоп-кодон). Таким образом, если в третьем положении находятся C или U, то кодон не является уникальным. Аденин в третьем положении не встречается у кодонов, соответствующих единственной аминокислоте.

Согласно существующим правилам написания последовательностей нуклеиновых кислот, последовательность антикодона (как и любую другую последовательность) всегда записывают в направлении от 5′-конца к

Третье основание кодона имеет наименьшее значение для его специфичности			
UUU UUC	UCU UCC	UAU UAC	UGU UGC
UUA UUG	UCA UCG	UAA UAG	UGA UGG
CUU CUC CUA CUG	CCU CCC CCA CCG	CAU CAC	CGU CGC CGA CGG
		CAA CAG	
AUU AUC AUA	ACU ACC ACA ACG	AAU AAC	AGU AGC
AUG		AAA AAG	AGA AGG
GUU GUC GUA GUG	GCU GCC GCA GCG	GAU GAC	GGU GGC GGA GGG
		GAA GAG	

Эффект, оказываемый третьим основанием	Основания в третьем положении, не меняющие значения кодона	Число кодонов
третья позиция в кодоне	U, C, A, G U, C, A	32 3
Не играет роли	A или G U или C	14 10
Придает кодону уникальное значение	Только G	2
терминирующие кодоны (прим.)		

РИС. 9.3 Третье основание кодона наименее существенно для его специфичности. Выделенные цветом рамки показывают группы кодонов, имеющих одно и то же значение вследствие вырожденности третьего основания

3′-концу. Поэтому, несмотря на то что антикодон комплементарен кодону и третье основание кодона взаимодействует с первым основанием антикодона, комбинацию

Кодон	5′	ACG	3′
Антикодон	3′	UGC	5′

обычно записывают как кодон ACG/антикодон CGU (вследствие правила записи последовательностей антикодон написан как бы наоборот, и читать его надо справа налево).

Во избежание путаницы, мы будем записывать все последовательности как обычно, в направлении от 5′ к 3′, но при этом последовательность антикодона будем обозначать стрелкой, направленной в обратную сторону (для напоминания об имеющейся взаимосвязи с кодоном). Так, рассмотренная выше пара кодон/антикодон будет записана как ACG и $\overleftarrow{\text{CGU}}$ соответственно.

Существует ли для каждого кодона своя тРНК с полностью комплементарным антикодоном? Или одна тРНК может взаимодействовать со всеми (или, по крайней мере, с несколькими) представителями кодонового семейства?

Показано, что одна тРНК может узнавать более одного кодона. Следовательно, основание в первом положении антикодона должно обладать способностью образовывать пару с несколькими различными основаниями, находящимися в третьем положении соответствующих кодонов. Как уже было сказано, взаимодействие оснований в первом положении антикодона не ограничено образованием канонических пар G-C и A-U.

Правила, описывающие кодон-антикодоновые взаимодействия, суммированы в **гипотезе неоднозначного спаривания (гипотеза качелей,** wobble hypothesis). Эта гипотеза гласит, что образование пары кодон-антикодон в двух первых положениях кодона всегда происходит по каноническим правилам, но в третьем положении возможно «колебание» (неоднозначное спаривание, или воблинг). Объясняется это тем, что конформация антикодоновой петли тРНК допускает значительную подвижность первого основания антикодона. На РИС. 9.4 приведен пример образования

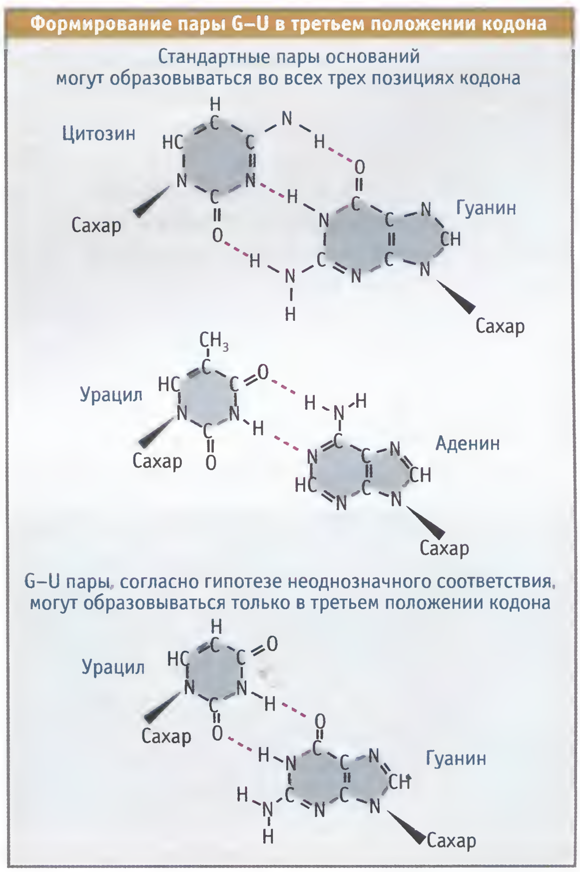


РИС. 9.4 Образование пары G-U между третьим основанием кодона и первым основанием антикодона согласно гипотезе неоднозначного соответствия

Воблинг третьего основания кодона	
Основание в первой позиции антикодона	Основание в третьей позиции кодона, с которыми возможно образование связи
U	A или G
C	только G
A	только U
G	C или U

РИС. 9.5 В третьей позиции кодон-антикодоновой пары возможно неоднозначное спаривание

пары G-U в дополнение к обычным взаимодействиям. Возникает новая модель взаимодействия, в которой A больше не имеет однозначного соответствия (так как U, обычно комплементарный аденину, должен узнавать и G). По аналогии, C также не имеет однозначного соответствия (так как обычно комплементарный цитозину гуанин должен узнавать еще и U). На **РИС. 9.5** приведены возможные варианты взаимодействия.

Исходя из вышесказанного можно сделать вывод, что однозначное узнавание кодона возможно, только если в третьем положении стоит G или U. Однако и это правило используется не часто. Примерами таких кодонов могут служить UGG и AUG, однако аналогичных кодонов с третьим U не существует. (G-U пары вообще свойственны для двухцепочечных РНК, однако для случая кодон-антикодонового взаимодействия, которое фиксируется только спариванием оснований, формирование стабильных взаимодействий особенно актуально, и G-U пары могут вносить вклад в это формирование, лишь находясь в последней позиции кодона.)

9.4 тРНК образуется из длинного предшественника

Ключевые понятия:

- Зрелая тРНК образуется в результате процессинга предшественника.
- 5′-конец зрелой тРНК образуется в результате расщепления молекулы-предшественника эндонуклеазой — РНКазой P.
- 3′-конец образуется в результате расщепления, которое следует за удалением нескольких последних оснований и добавлением общего для всех тРНК терминального триплета CCA.

Как правило, тРНК синтезируется в виде предшественника, имеющего по краям дополнительные последовательности. На **РИС. 9.6** показано, как «избыточные» последовательности удаляются комбинацией ферментов с экзо- и эндонуклеазными активностями. Общая для всех зрелых тРНК 3′-концевая последовательность CCA не кодируется геномом, а добавляется к структуре тРНК в результате процессинга.

5′-конец тРНК формируется после отщепления «излишков» при участии фермента — рибонуклеазы P.

Ферменты, осуществляющие процессинг 3′-конца тРНК, лучше всего охарактеризованы у *E. coli*. Эндонуклеазы запускают процессинг молекулы-предшественника, а затем экзонуклеазы отщепляют избыточные концевые нуклеотиды по направлению от 3′- к 5′-концу. У эукариот в этот процесс вовлечено несколько ферментов. В результате их работы образуется молекула тРНК, к 3′-концу которой на завершающем этапе процессинга должна быть присоединена последовательность CCA.

Концевая последовательность CCA не кодируется генами тРНК, ее добавление является результатом специального ферментативного процесса. Специфичность присоединения CCA определяется белком,

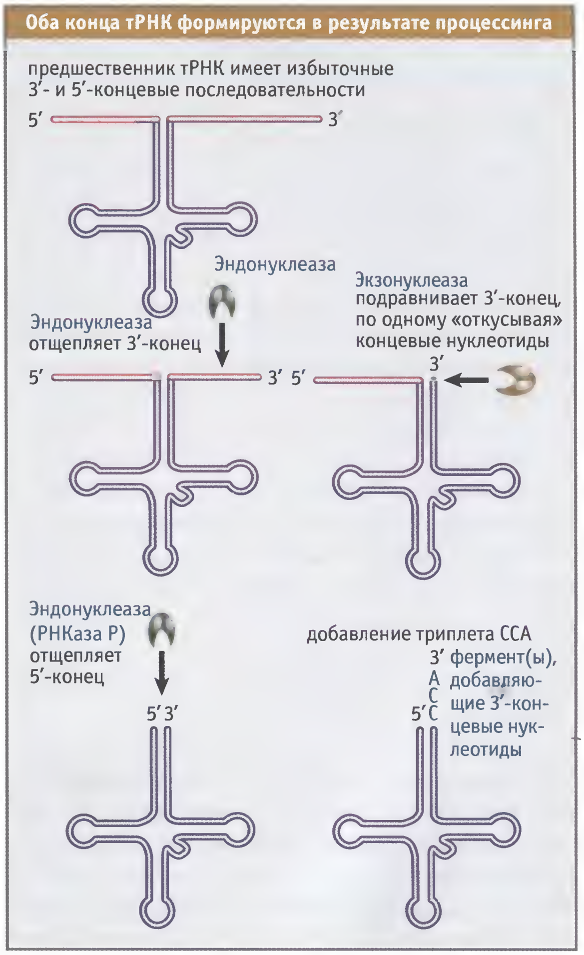


РИС. 9.6 3′-конец тРНК образуется в результате отщепления избыточных концевых нуклеотидов и последующего добавления последовательности CCA; 5′-конец образуется в результате отщепления концевой последовательности

добавляющим эти три нуклеотида к любой последовательности тРНК, где она отсутствует. Существует несколько вариантов этого процесса, отличающихся у разных организмов.

У некоторых организмов присоединение терминальной последовательности ССА катализируется одним единственным ферментом. По одной из версий этот фермент связывает 3′-конец тРНК и поочередно добавляет к нему нуклеотиды С, С и А. Специфичность каждой стадии определяется структурой 3′-конца молекулы тРНК. Другая модель предполагает, что фермент имеет разные активные сайты связывания для цитозинтрифосфата (ЦТФ) и аденозинтрифосфата (АТФ).

В других организмах за добавление С и А — нуклеотидных остатков отвечают разные ферменты, которые действуют последовательно.

В случае ошибок процессинга тРНК активируется «система контроля качества», запускающая процесс деградации бракованной молекулы. Это исключает блокировку аппарата белкового синтеза нефункциональными тРНК.

9.5 тРНК содержит модифицированные основания

Ключевые понятия:

- тРНК содержит более 50 модифицированных оснований.
- Модификация обычно представляет собой изменение исходных нуклеотидов тРНК, однако в некоторых случаях заключается в замене одних оснований на другие.

Одна из характерных особенностей транспортных РНК, отличающая их от других нуклеиновых кислот, — это наличие в составе молекулы «необычных» оснований. Необычным является любое пуриновое или пиримидиновое основание, отличное от A, G, C или U, из которых изначально синтезируются все РНК. Все нестандартные основания образуются в результате **модификаций** одного из четырех вышеназванных оснований (но уже после того, как основание будет включено в РНК).

Все классы РНК демонстрируют некоторую степень модификации оснований, но во всех случаях, кроме тРНК, эти модификации сводятся к простым реакциям (таким, как метилирование оснований). В случае же с тРНК можно увидеть разнообразные модификации, начиная от простого метилирования и до крупномасштабных перестроек пуриновых колец. Модификации затрагивают все части молекулы тРНК. Описано более 50 различных модифицированных оснований, встречающихся в тРНК.

На **РИС. 9.7** изображены некоторые из наиболее часто встречающихся в тРНК модифицированных оснований. Модификации пиримидинов (C и U) менее сложны, чем изменения структуры пуринов (A и G). Кроме модификаций, затрагивающих структуру самого азотистого основания, можно обнаружить изменения, связанные с метилированием рибозного кольца в 2′–O положении.

Наиболее часто встречаются модификации уридина. Метилирование в пятом положении кольца приводит к появлению риботимидина (T). Такое же основание обнаруживается и в ДНК, но если в ДНК тимидин соединен с дезоксирибозой, то в случае тРНК он связан с рибозой. Присутствие в РНК тимина является необычным для данного типа молекул.

Дигидроуридин (D) образуется в результате насыщения двойной связи, изменяющей структуру кольца. В псевдоуридине (ψ) атомы N и C меняются местами (см. рис. 26.40). В 4-тиоуридине на месте атома кислорода находится атом серы.

Инозин — нуклеозид, который обычно обнаруживается в клетке в качестве промежуточного продукта при биосинтезе пуринов. Тем не менее он не включается непосредственно в РНК, но превращение A в I необходимо для ее формирования. Другие модификации A включают присоединение различных сложных групп.

При модификации G образуются две сложные серии нуклеотидов. Основания Q-типа, такие как квеуозин, содержат дополнительное пентозное кольцо, присоединенное с помощью NH-связи к метильной группе 7′-метилгуанозина. Пентозное кольцо, в свою очередь, может содержать различные дополнительные группы. Основания Y-типа, такие как виозин, содержат дополнительное кольцо, слитое воедино с пуриновым кольцом. Это внешнее кольцо несет длинную углеродную цепь, к которой могут быть присоединены различные дополнительные группы.

Реакция модификации обычно заключается в изменении (или добавлении) определенной группы к основаниям, уже включенным в состав тРНК. Исключением является синтез оснований Q-типа, для которых был обнаружен фермент, способный обменивать свободный квеуозин на остаток гуанина в тРНК. Реакция состоит из разрыва и последующего восстановления связей по обе стороны от нуклеозида.

Модифицированные нуклеозиды синтезируются с помощью особых РНК-модифицирующих ферментов. Нуклеозид, изначально присутствующий в каждом положении тРНК, может быть идентифицирован путем сравнения последовательности тРНК и кодирующего ее гена или (что более трудоемко) путем выделения молекул-предшественников тРНК, не подвергшихся модификациям. С помощью последнего подхода было показано, что различные модификации вносятся в тРНК на разных стадиях ее образования из предшественника.

Ряд модифицированных оснований присутствует во всех молекулах тРНК — например, D-остаток, давший название D-петле, и ψ, находящийся в TψU —

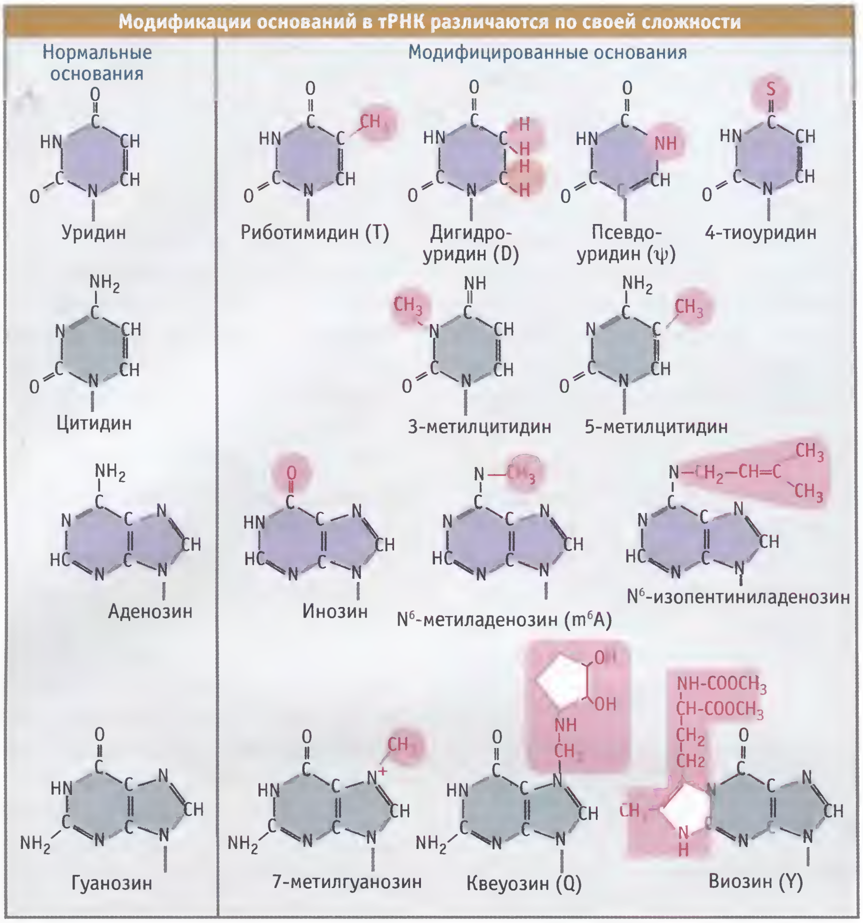


РИС. 9.7 Любое из четырех оснований тРНК может быть модифицировано

последовательности. С 3′-конца антикодона всегда находится модифицированный пурин, хотя характер этой модификации может варьировать в широких пределах.

Другие модификации специфичны для определенных тРНК или групп тРНК. Например, присутствие виозина характерно для $тРНК^{Phe}$, выделяемой из бактерий, дрожжей и млекопитающих.

Многочисленные тРНК-модифицирующие ферменты (~60 различных белков у дрожжей) широко варьируют по своей специфичности. В одних случаях ферменты могут определенным образом модифицировать основания только в одном положении. В других ферменты могут модифицировать основания, находящиеся в нескольких разных положениях. Некоторые ферменты катализируют строго определенные реакции с индивидуальными тРНК; другие обладают широкой субстратной специфичностью. Каким образом тРНК-модифицирующие ферменты узнают свои субстраты, неизвестно. Возможно, узнавание обеспечивается структурными особенностями молекулы тРНК в районе сайта модификации. Некоторые модификации требуют для своего осуществления участия нескольких ферментов.

9.6 Модифицированные основания влияют на кодон-антикодоновые взаимодействия

Ключевые понятия:

- Модификации нуклеотидов антикодона влияют на набор воблирующих оснований, тем самым определяя специфичность тРНК.

Наиболее ощутимый эффект на свойства тРНК оказывают модификации оснований антикодона, что может вести даже к изменению значения тРНК за счет влияния на способность антикодона взаимодействовать с кодоном. Модификации в участках молекулы тРНК, расположенных по соседству с антикодоном, также влияют на его взаимодействие с кодоном.

Если основания антикодона модифицированы, то возможные варианты взаимодействия уже не ограничиваются четырьмя стандартными основаниями (A, G и U, C), образующими комплементарные и воблирующие пары. На **РИС. 9.8** показаны варианты взаимодействия инозина (I), часто присутствующего в первом

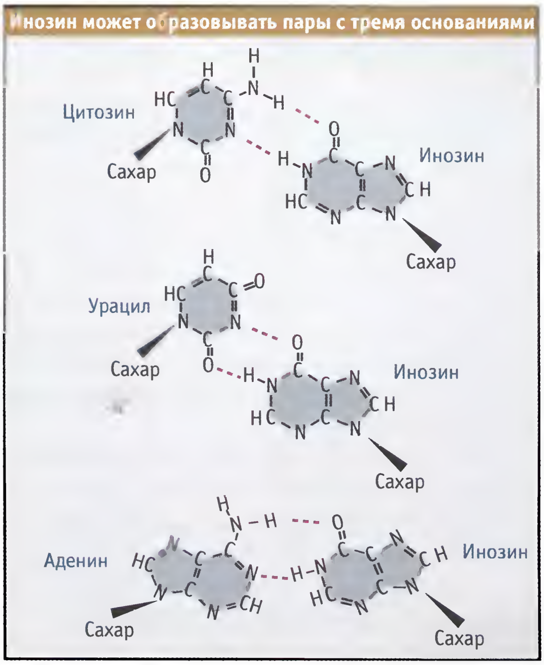


РИС. 9.8 Инозин может образовывать пары с U, C и A

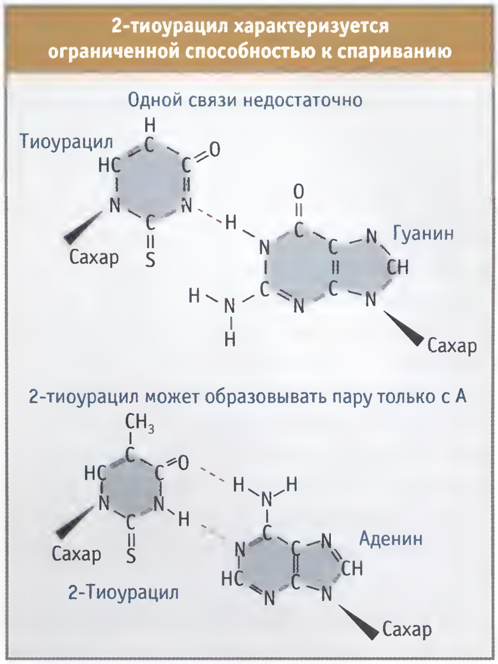


РИС. 9.9 В результате модификации 2-тиоуридин теряет способность образовывать связи с G и может формировать их только с A

положении антикодона, с тремя основаниями — U, C и A.

Это особенно важно для изолейциновых кодонов, где AUA кодирует изолейцин, в то время как AUG кодирует метионин. Когда A находится в третьем положении кодона, при помощи обычных оснований невозможно образовать однозначную пару, так как любая тРНК данной группы кодонов, у которой в первом положении антикодона находится U, будет воспринимать AUG и AUA как равнозначные. Однако в соответствии с генетическим кодом триплет AUA должен читаться так же, только как AUU и AUC. Это возможно благодаря существованию изолейциновых тРНК, содержащих инозин в первом положении антикодона.

В действительности, некоторые классические варианты пар никогда не встречаются при кодон-антикодоновом взаимодействии в силу того, что отдельные основания всегда модифицированы. Так, в первом положении антикодона никогда не обнаруживается A, поскольку обычно он превращается в I. В большинстве случаев U в первом положении антикодона превращается в модифицированную форму с измененной способностью к образованию пар.

В результате некоторых модификаций возникают основания, также влияющие на избирательность узнавания кодонов. Антикодоны, несущие в первом положении уридин-5-оксиуксусную кислоту или 5-метоксиуридин, эффективно узнают A и G в третьем положении кодона и менее эффективно узнают U. Другой пример неоднозначного спаривания (но с существованием более предпочтительного варианта) обнаруживается в ряду квеуозина и его производных. Этот модифицированный гуанин сохраняет способность узнавать как C так и U, но предпочтительно взаимодействует с U.

Ограничение образования стандартных пар возможно при наличии в антикодоне 2-тиоуридина. На **РИС. 9.9** показано, что эта модификация не нарушает способности тимина образовывать пару с A, но запрещает его взаимодействие при воблинг-спаривании с G.

Рассмотрение различных вариантов кодон-антикодоновых взаимодействий позволяет сделать вывод о том, что существует несколько способов для создания набора молекул тРНК, способных узнавать 61 смысловой кодон. В любом конкретном организме ни один из способов не имеет преимуществ, хотя отсутствие определенного пути модификации может помешать использованию некоторых способов узнавания. Таким образом, в разных организмах одно и то же семейство кодонов может «считываться» разными тРНК.

В клетке часто присутствуют тРНК, чьи способности к узнаванию кодонов перекрываются. Вследствие этого определенный кодон может быть прочитан более чем одной тРНК. При этом возможны различия в эффективности связывания кодонов. (Как правило, кодоны, встречающиеся чаще, читаются более эффективно.) С другой стороны, в составе полного набора тРНК (необходимого для считывания всех кодонов) присутствует множество

молекул, способных взаимодействовать с несколькими разными кодонами.

Почти для всех тРНК предсказания гипотезы неоднозначного спаривания хорошо согласуются с имеющимися данными. Но существует ряд исключений, при которых кодон-антикодоновое взаимодействие описывается механизмами, не предусмотренными правилами гипотезы. Вероятно, эти эффекты обусловлены влиянием соседних оснований и/или конформационными изменениями в антикодоновой петле. Действительно, сама идея, заложенная в гипотезу неоднозначного соответствия, постулирует важность структуры антикодоновой петли. Подтверждением влияния ближайшего окружения антикодона на его способность узнавать кодоны можно считать результаты изучения спонтанных мутантов, несущих замены вне антикодоновой области молекулы. Эти мутации, несмотря на их локализацию, меняют взаимодействие тРНК с кодоном.

Другим примером взаимодействия, не предсказанного гипотезой, является способность бактериальной инициаторной fMet-тРНК$_f$ узнавать как кодон AUG, так и кодон GUG. Это вынуждает третье основание антикодона взаимодействовать «не по правилам».

9.7 Изменения в универсальном генетическом коде носят спорадический характер

Ключевые понятия:

- У некоторых видов встречаются отклонения от универсального генетического кода.
- Изменения в генетическом коде наиболее характерны для генома митохондрий, на основании этих изменений можно построить филогенетическое древо.
- В ядерном геноме отклонения от стандарта крайне редки и обычно затрагивают лишь стоп-кодоны.

Генетический код характеризуется универсальностью, однако из этого правила существуют некоторые исключения. Исключения эти, как правило, затрагивают кодоны, вовлеченные в инициацию или терминацию белкового синтеза, и появляются в результате возникновения (или утраты) тРНК, способных узнавать такие кодоны. Отклонения от «классики», обнаруженные в бактериальных и эукариотических геномах, суммированы на **РИС. 9.10**.

Почти все изменения, в результате которых кодон приобретает способность кодировать новую аминокислоту, касаются стоп-кодонов:

- В клетках бактерии *Mycoplasma capricolum* кодон UGA используется не для терминации, а для кодирования триптофана (Trp). Причем использование UGA в качестве Trp-кодона предпочтительней, чем использование UGG. Существует два вида Trp-тРНК: с антикодоном UCA← (прочитывает кодоны UGA и UGG) и с антикодоном CCA← (прочитывает только UGG).
- У некоторых простейших из типа ресничных кодоны UAA и UAG используются не в качестве терминирующего сигнала, а для кодирования глутамина. Так, у *Tetrahymena termophila* имеется три вида тРНК^{Glu}. Один из них узнает обычные кодоны для глутамина — CAA и CAG, второй узнает UAA и UAG (согласно гипотезе неоднозначного спаривания), а третий — только UAG. Мы можем допустить, что дополнительные изменения заключаются в том, что фактор терминации eRF приобретает ограниченную специфичность в сравнении с другими эукариотами.
- У других ресничных (*Euplotes octacarinatus*) кодон UGA кодирует цистеин. В качестве стоп-кодона используется только кодон UAA, а кодон UAG не обнаружен. Изменения в значении кодона UGA дополняются модификацией антикодона тРНК^{Cys}, которая позволяет читать и кодоны UGU и UGC.
- Смена значения кодона с одной аминокислоты на другую встречается у дрожжей (*Candida*), у которых CUG кодирует серин вместо лейцина (а UAG используется в качестве смыслового кодона).

Приобретение стоп-кодонами кодирующей функции опосредуется изменениями двух типов: тРНК должна мутировать таким образом, чтобы узнавать стоп-кодон, а фактор терминации должен мутировать таким образом, чтобы не распознавать кодон как терминирующий.

Другим часто встречающимся вариантом изменения генетического кода является утрата тРНК, отвечающей за тот или иной кодон, и, как следствие, неспособность кодона транслироваться в аминокислоту. Дальнейшая судьба такого кодона зависит от того, будет ли он распознаваться фактором терминации трансляции.

Все описанные изменения генетического кода единичны. Это позволяет утверждать, что возникали они в ходе эволюции независимо друг от друга. Как правило, изменения выражаются в замене значения стоп-кодонов, а не в замене кодона одной аминокислоты на кодон другой. После того как на ранних этапах эволюции сложилась система кодирования, любое изменение значения кодона вело к изменению аминокислоты во всех белках, содержащих данную аминокислоту. По крайней мере, для части белков такие изменения могли стать критичными, что потребовало бы начинать процесс селекции белков заново. Двоякое использование терминирующих кодонов отражает возможность их вовлечения в процессы информационного кодирования. Если некоторые из кодонов терминации встречаются очень редко, они могут быть использованы для кодирования аминокислот в результате изменений, позволяющих тРНК узнавать их.

Исключения из универсального генетического кода часто обнаруживаются в митохондриях. На **РИС. 9.11** показано филогенетическое древо таких изменений. Оно предполагает, что сначала существовал универсальный код, который претерпевал изменения на разных этапах

Изменения генетического кода обычно затрагивают стоп-кодоны либо приводят к потере кодоном своего значения

UUU Phe UUC UUA Leu UUG	UCU UCC Ser UCA UCG	UAU Tyr UAC UAA СТОП→Gln UAG	UGU Cys UGC UGA СТОП→Trp, Cys, Sel UGG Trp
CUU CUC Leu CUA CUG Leu→Ser	CCU CCC Pro CCA CCG	CAU His CAC CAA Gln CAG	CGU CGC CGA Arg CGG Arg→УТРАТА
AUU AUC Ile AUA Ile→УТРАТА AUG Met	ACU ACC Thr ACA ACG	AAU Asn AAC AAA Lys AAG	AGU Ser AGC AGA Arg→УТРАТА AGG Arg
GUU GUC Val GUA GUG	GCU GCC Ala GCA GCG	GAU Asp GAC GAA Glu GAG	GGU GGC GGA Gly GGG

РИС. 9.10 Изменения генетического кода в геномах бактериальных и эукариотических ядер приводят либо к утрате кодоном способности кодировать аминокислоту, либо придают терминирующим кодонам смысловое значение. Изменения значения кодона с одной аминокислоты на другую, как правило, не происходит

эволюции митохондрий. Согласно рисунку, самым первым изменением, характерным для всех митохондрий (кроме растительных), стало использование UGA для кодирования триптофана.

Некоторые из изменений привели к упрощению кода, через замещение пары кодонов, имеющих разное значение, парой с одинаковым значением. Примерами могут служить кодоны UGG и UGA, кодирующие Trp (вместо одного кодона для Trp и одного для терминации), кодоны AUG и AUA, кодирующие Met (вместо одного кодона для Met и одного для Ile).

Почему возникли изменения в коде митохондрий? Система митохондриального белкового синтеза, как правило, производит небольшое число разных белков (всего около 10). Поэтому проблема значительных внутриклеточных нарушений из-за смены значения кодонов могла быть не столь остра. Вероятно, измененные кодоны располагались в таких положениях, в которых аминокислотные замены не приводили к серьезным повреждениям. Разнообразие изменений генетического кода, найденных в митохондриях различных видов, ставит под сомнение возможность их происхождения от общего предкового митохондриального кода и позволяет предположить, что они возникали и эволюционировали независимо друг от друга.

Согласно гипотезе неоднозначного соответствия, требуется как минимум 31 тРНК (без учета специальных инициаторных тРНК) для того, чтобы распознать 61 кодон (по крайней мере две тРНК необходимы для каждого семейства кодонов, и одна тРНК — для кодоновой пары или единичного кодона). Но в некоторых митохондриях (например, у млекопитающих) наблюдается необычная ситуация: они содержат только 22 разных



РИС. 9.11 Для изменений генетического кода генома митохондрий может быть сконструировано филогенетическое древо. Предположительно, замены, превратившие AUA в кодон, кодирующий Met, а AAA –в кодон, кодирующий Asn, произошли в эволюции дважды и независимо, а замена, вернувшая кодону AUA его исходное значение, произошла у иглокожих

тРНК. Как же этот неполный набор тРНК обслуживает все кодоны?

Ключевым моментом, лежащим в основе этого явления, стало упрощение кодон-антикодонового взаимодействия, в результате которого одна тРНК приобретает способность узнавать все четыре представителя кодонового семейства. Это позволяет свести минимальное число тРНК, необходимых для узнавания всех кодонов, к 23. Использование AG^A_G для терминации трансляции позволяет сократить требуемое число тРНК до 22.

У тРНК, обслуживающих все восемь кодоновых семейств, в первом положении антикодона находится немодифицированный U. Оставшиеся кодоны распределяются по парам: заканчивающиеся пиримидинами кодоны прочитываются основанием G (в первом положении антикодона), а все триплеты, заканчивающиеся пуринами, прочитываются модифицированным U в антикодоне (согласно правилам гипотезы неоднозначного спаривания). Сложность с уникальным кодоном UGG преодолевается в результате изменений в генетическом коде: кодон UGA прочитывается наряду с UGG как триптофан. AUA в митохондриях млекопитающих больше не кодирует изолейцин, а узнается наряду с AUG как метионин. В результате все кодоны, не вошедшие в семейства, подразделяются на 14 пар по способности взаимодействовать с молекулами тРНК.

22 идентифицированных митохондриальных гена тРНК кодируют 14 молекул тРНК, узнающих пары кодонов, и 8 молекул тРНК, узнающих семейства кодонов. Остаются два обычных стоп-кодона UAG и UAA (не узнаваемых тРНК), а также кодоновая пара AG^{A}_{G}. Сходные результаты получены при исследовании митохондрий грибов.

9.8 Некоторые стоп-кодоны могут встраивать в белок нестандартные аминокислоты

Ключевые понятия:

- Ряд генов санкционирует особое прочтение специфических кодонов.
- Взаимодействие селено-Cys-тРНК и UGA-кодона требует участия нескольких белков, модифицирующих Cys-тРНК и присоединяющих ее к рибосоме.
- Пирролизин может транслироваться с UAG-кодона.

В отдельных генах могут появляться особенные редкие изменения считывания кодонов. Специфичность таких изменений предполагает, что узнавание отдельных кодонов определяется их окружением. Известны два случая, когда аминокислоты, не входящие в список 20 канонических, встраиваются в белковую последовательность специальными аминоацил-тРНК.

Ярким примером является вставка в белковую последовательность модифицированной аминокислоты селеноцистеина в ответ на определенный кодон UGA гена, кодирующего селенопротеин у прокариот и эукариот. Обычно эти ферменты катализируют окислительно-восстановительные реакции и содержат один селеноцистеиновый остаток, формирующий часть активного центра белка. Наиболее изученным является процесс использования кодона UGA у трех генов *E.coli*, кодирующих изоформы формитдегидрогеназы. Внутренний кодон UGA в этих случаях читается селено-Cys-тРНК. Эта необычная реакция опосредуется локальной вторичной структурой молекулы мРНК, в частности шпилькой, расположенной за UGA кодоном.

Мутации в четырех генах *sel* ведут к образованию систем синтеза, дефицитных по селенопротеинам. *selC* кодирует тРНК (с антикодоном $ACU^{\leftarrow}$), соответствующую аминокислоте серин. Продукты генов *selA* и *selD* необходимы для того, чтобы превратить серин в селеноцистеин. Белок SelB представляет собой альтернативный фактор элонгации, это гуанин-связывающий фермент, выступающий в качестве специального трансляционного фактора, обеспечивающего попадание селено-Cys-тРНК в A-сайт рибосомы, и являющийся аналогом фактора EF-Tu (для селено-Cys-тРНК). Последовательность SelB гомологична как EF-Tu, так и IF-2.

Почему же селено-Cys-тРНК вставляется только по определенному UGA кодону? За этим кодоном следует участок мРНК с особой вторичной структурой (типа стебель–петля). На **РИС. 9.12** показано, как стебель этой структуры узнается дополнительным доменом белка SelB (отсутствующим у фактора EF-Tu и фактора IF-2). С помощью похожего механизма интерпретируются некоторые кодоны UGA в клетках млекопитающих [за исключением того, что для узнавания соответствующих UGA необходимо сразу два белка: один белок (SBP2) связывает структуру типа стебель-петля, расположенную за кодоном UGA, в то время как аналог SelB (называемый SECIS) одновременно связывает SPB2 и соответствующую тРНК].

Другим примером включения особых аминокислот в последовательность белка является вставка пирролизина в ответ на кодон UAG. Это явление встречается у архей и бактерий. Механизм похож на механизм вставки селеноцистеина. Необычная тРНК связывает лизин, который лишь затем претерпевает модификацию. Эта тРНК имеет антикодон CUA, взаимодействующий с кодоном UAG. В процессе должны участвовать и другие компоненты, ограничивающие взаимодействие нестандартной тРНК с «обычными» UAG.

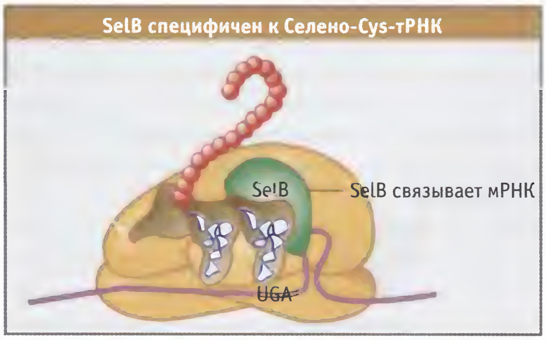


РИС. 9.12 Белок SelB — фактор элонгации, специфично связывающий селено-Cys-тРНК с кодоном UGA, за которым следует участок сложной вторичной структуры мРНК (стебель–петля)

9.9 Синтетазы присоединяют аминокислоты к тРНК

Ключевые понятия:

- Аминоацил-тРНК-синтетазы — ферменты, присоединяющие аминокислоты к молекулам тРНК. Образование аминоацил-тРНК идет в 2 стадии с использованием энергии АТФ.
- В каждой клетке присутствует 20 аминоацил-тРНК-синтетаз. Каждой аминокислоте соответствует как минимум одна синтетаза, которая обнаруживает все тРНК, представляющие эту аминокислоту.
- Узнавание молекулой тРНК своей аминокислоты обеспечивается небольшим числом сайтов связывания в ее нуклеотидной последовательности.

Все тРНК должны быть одновременно схожи по одним характеристикам и быть различимыми по другим. Это требование реализуется благодаря способности молекул тРНК сворачиваться определенным образом, формулируя при этом характерную пространственную структуру (похожую на бумеранг или букву L). Вариабельность отдельных деталей этой структуры, таких как угол между двумя плечами L-образной структуры или выпячивание отдельных оснований, отличает тРНК друг от друга.

Все тРНК обладают способностью связываться с A и P сайтами рибосомы. При этом одним концом они взаимодействуют с кодоном мРНК, а другой конец участвует в присоединении доставленной аминокислоты к растущему полипептиду. Кроме того, все тРНК (за исключением специальных инициаторных тРНК) обладают способностью взаимодействовать с факторами элонгации (EF-Tu или eEF1), участвующими в связывании тРНК с рибосомой. Инициаторные же тРНК узнаются собственными факторами инициации IF2 и eIF2. Следовательно, для взаимодействия с факторами элонгации молекулы тРНК должны обладать определенными общими чертами (при этом отличающими их от инициаторных тРНК).

Аминокислоты вовлекаются в белковый синтез посредством аминоацил-тРНК-синтетаз. Эти ферменты обеспечивают правильное взаимодействие аминокислот с тРНК. Все синтетазы функционируют согласно двухступенчатому механизму, приведенному на РИС. 9.13:

- Сначала аминокислота взаимодействует с АТФ с образованием аминоацил-аденилата (при этом высвобождается пирофосфат). Необходимая для реакции энергия образуется в результате расщепления макроэргических связей АТФ.
- Затем активированная аминокислота переносится на тРНК с высвобождением АМФ.

Синтетазы подбирают тРНК и аминокислоты в строгом соответствии с генетическим кодом. Каждая синтетаза узнает только одну аминокислоту и все типы тРНК, соответствующие данной аминокислоте (обычно одной аминокислоте соответствует более одной тРНК). Для узнавания синонимичных кодонов может быть необходима более чем одна тРНК. Иногда в клетке присутствует несколько видов тРНК, реагирующих с одним и тем же кодоном. Молекулы тРНК, соответствующие одной аминокислоте, называют **изоакцепторными.** А так как все они узнаются одной и той же синтетазой, то имеют и второе название — **родственные тРНК.**

Многочисленные попытки выявить сходство последовательностей изоакцепторных тРНК, а также опыты по индукции химических модификаций, влияющих на специфичность присоединения аминокислот, показали, что механизмы распознавания разных тРНК аминоацил-тРНК-синтетазой различаются и не ограничиваются какой-либо одной особенностью первичной или вторичной структуры молекулы. Анализ кристаллов тРНК показал, что акцепторный и антикодоновый стебли образуют тесный контакт с синтетазой. Мутации, способные влиять на специфичность взаимодействия тРНК с синтетазой, локализованы именно в этих двух регионах. (При этом изменения в самом антикодоне необязательно ска-

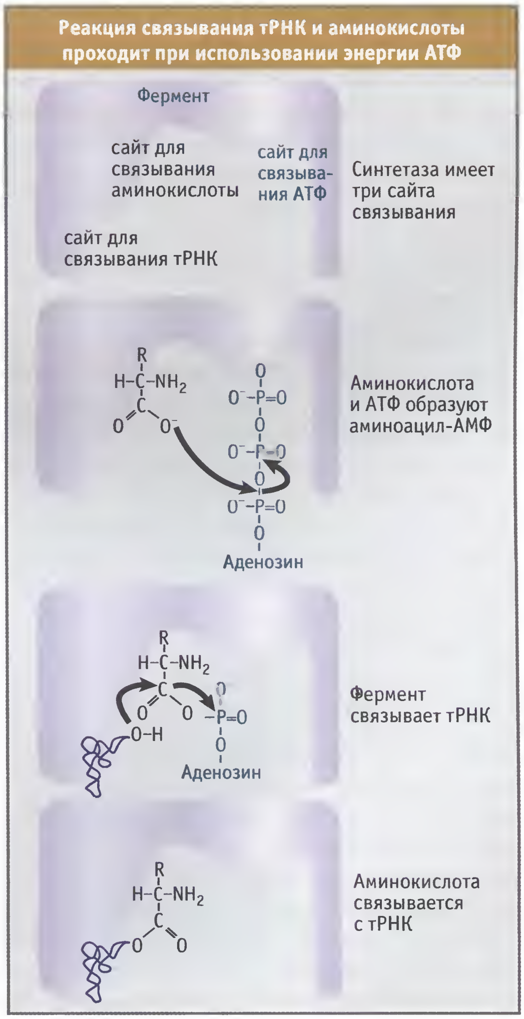


РИС. 9.13 Аминоацил-тРНК-синтетазы присоединяют аминокислоты к молекулам тРНК

зываются на способности тРНК связываться с ферментом; к примеру, рассматриваемые далее в этой главе супрессорные мутации изменяют основание в антикодоне, а следовательно, и кодон, с которым взаимодействует тРНК, но не влияют на тип аминокислоты, присоединяемой к тРНК.)

Группа изоакцепторных тРНК должна взаимодействовать с одной синтетазой, специфичной для соответствующей аминокислоты, а потому должна обладать и общими свойствами, позволяющими ферменту отличать тРНК этой группы от всех других тРНК. Все тРНК делятся на 20 изоакцепторных групп, и каждая группа узнается своей аминоацил-тРНК-синтетазой.

тРНК взаимодействует с синтетазой при помощи участков, состоящих из небольшого числа оснований (как правило, от одного до пяти). Взаимодействие тРНК и фермента характеризуется тремя основными чертами:

- Как правило (но не всегда), во взаимодействии участвует одно основание антикодона (но иногда все три положения антикодона играют роль в узнавании).
- Часто во взаимодействии участвует одна из трех последних пар оснований в акцепторной петле. Например, аланиновая тРНК узнается ферментом по единственной уникальной паре оснований в акцепторной петле.
- Между акцепторной петлей и 3′-концевой последовательностью CCA находятся так называемые дискриминирующие основания, которые всегда неизменны в пределах одной группы изоакцепторных тРНК.

Ни одна из этих особенностей сама по себе не обеспечивает достаточной специфичности и не является определяющей при разделении всех тРНК на 20 групп. Узнавание тРНК своим ферментом — уникальный процесс, подчиняющийся в каждом конкретном случае своим правилам.

Некоторые синтетазы узнают тРНК по «мини-спирали», состоящей из акцепторного участка молекулы тРНК и TψC-плеча (что соответствует одному плечу L-образной молекулы тРНК). Специфичность узнавания ферментами некоторых тРНК зависит исключительно от структуры их акцепторного стебля. Тем не менее очевидно, что существуют значительные вариации между разными тРНК, и в ряде случаев важную роль играет и антикодоновая петля. Мутации в антикодоне оказывают влияние на узнавание тРНК Phe-тРНК-синтетазой II класса. В процесс узнавания может быть вовлечено сразу несколько факторов: для правильного узнавания синтетазами $тРНК^{Val}$ и $тРНК^{Met}$ *in vivo* важны одновременно антикодон и миниспирали.

Таким образом, узнавание определяется взаимодействием между несколькими контактными точками в тРНК (большей частью расположенными по концам молекулы) и несколькими аминокислотами, формирующими субстратсвязывающий сайт фермента. Отношение вкладов акцепторного и антикодонового участков молекулы тРНК в это взаимодействие варьирует для разных пар тРНК-синтетаза.

9.10 Аминоацил-тРНК-синтетазы делят на 2 класса

Ключевые понятия:

- В зависимости от аминокислотной последовательности и структурного сходства аминоацил-тРНК-синтетазы делят на 2 класса.

Несмотря на сходство функций, аминоацил-тРНК-синтетазы представляют собой довольно гетерогенную группу белков. Размеры отдельных субъединиц могут варьировать от 40 до 110 кДа, а сами ферменты могут существовать в виде мономеров, димеров или тетрамеров. Гомология между ними обнаруживается редко. Активный центр синтетазы представлен небольшой частью молекулы. При сравнении активных центров разных ферментов были получены довольно любопытные результаты.

На основе особенностей структуры домена, образующего активный центр фермента, все синтетазы разделены на два класса, каждый из которых состоит из десяти ферментов. Общий тип организации, свойственный обеим группам, приведен на РИС. 9.14. Каталитичесий домен, включающий сайты связывания АТФ и аминокислоты, представлен протяженным участком молекулы. Он прерывается вставкой домена, ответственного за взаимодействие с акцепторной спиралью тРНК, в результате чего акцепторный конец молекулы тРНК располагается в непосредственной близости от каталитического участка синтетазы. Отдельный домен отвечает за связывание антикодонового плеча тРНК. Синтетазы, существующие в виде мультимеров, часто имеют дополнительный олигомеризационный домен.

Синтетазы I класса обладают N-концевым каталитическим доменом, отличающимся наличием двух ко-

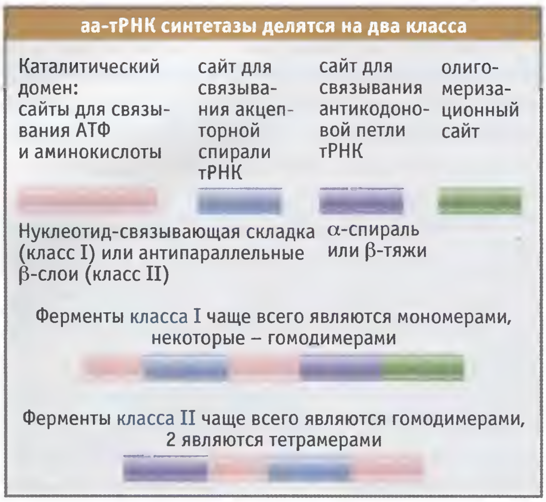


РИС. 9.14 Аминоацил-тРНК-синтетаза представлена тремя или четырьмя участками, каждый из которых выполняет свою функцию. (Только синтетазы, имеющие мультимерную структуру, обладают олигомеризационным доменом.)

ротких частично консервативных последовательностей, иногда называемых знаковыми последовательностями (asignature sequences). Каталитический домен формирует мотив, называемый нуклеотид-связывающей петлей (структуры подобного типа обнаруживаются и у других классов нуклеотид-связывающих ферментов). Данная петля состоит из перемежающихся β-слоев и α-спиралей; знаковые последовательности формируют часть АТФ-связывающего сайта. Вставки, контактирующие с акцепторной спиралью тРНК, отличаются у разных синтетаз I класса. Большим разнообразием характеризуются и С-концевые домены, включающие домен для связывания антикодонового стебля тРНК, а также олигомеризационные домены.

Для последовательностей каталитического домена синтетаз II класса характерны три общие особенности. Активный центр состоит из больших антипараллельных β-слоев, окруженных α-спиралями. Домен, ответственный за связывание акцепторной спирали тРНК, прерывает последовательность каталитического домена (по аналогии с синтетазами I класса), а его структура определяется особенностями каждого конкретного фермента. Антикодон-связывающий домен, как правило, является N-концевым. Локализация олигомеризационных доменов варьирует в широких пределах.

Отсутствие сколько-нибудь веских доказательств родства между двумя группами синтетаз — явление труднообъяснимое. Возможно, они возникли и эволюционировали независимо друг от друга. В этом случае можно предположить, что у ранних форм жизни белки состояли только из десяти аминокислот, взаимодействующих с ферментами одного из ныне существующих классов.

Согласно общепринятой модели, синтетаза связывает тРНК, контактируя с одной из сторон L-образной молекулы. Несколько общих принципов описывают все типы взаимодействий между аминоацил-тРНК-синтетазой и молекулой тРНК: связывание тРНК, главным образом, происходит по концам молекулы, при этом большая часть тРНК не участвует в процессе узнавания синтетазой. Однако детали процесса взаимодействия различаются для синтетаз I и II классов. Эти различия, выясненные с помощью данных рентгеноструктурного анализа, показаны на **РИС. 9.15**. Два класса ферментов связывают тРНК с противоположных сторон, в результате чего модели взаимодействия тРНК-белок для двух групп ферментов выглядят почти как зеркальные отражения друг друга.

Синтетаза I класса (Gln-тРНК-синтетаза)˙ контактирует с тРНК со стороны ее D-петли. Она узнает малую бороздку акцепторного стебля тРНК одним из краев своего контактного сайта и взаимодействует с антикодоновой петлей другим его краем. На **РИС. 9.16** схематично представлен комплекс тРНК^{Gln}-синтетаза. Контакт с ферментом изменяет структуру тРНК в двух точках. Эти изменения (в антикодоновой петле и акцепторном стебле) показаны с помощью пунктирных и сплошных линий:

- Основания U35 и U36 в антикодоновой петле достаточно далеко выступают за пределы тРНК и как бы погружаются в белок.
- Конец акцепторного стебля деформируется, в результате чего нарушается взаимодействие между U1 и A72. Одноцепочечный конец акцепторного стебля входит в глубокий карман в структуре фермента, который, кроме того, содержит сайт связывания АТФ.

Такая структура объясняет, почему изменения в основаниях U35, G73, а также паре U1-A72 влияют на процесс узнавания тРНК ферментом. Водородные связи возникают между тРНК и белком во всех этих положениях.

Синтетаза II класса (Asp-тРНК-синтетаза) контактирует с другой стороной тРНК; она узнает вариабельную

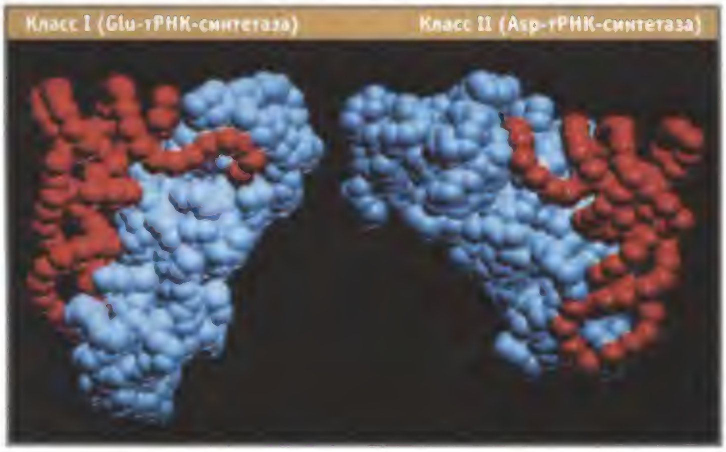


РИС. 9.15 Структура кристаллов аминоацил-тРНК-синтетаз I и II классов показывает, что ферменты разных классов связывают субстратную тРНК с противоположных сторон. тРНК изображена красным, белки — синим цветом. Любезно предоставлено Dino Moras, Institute of Genetics and Molecular and Cellular Biology

петлю и большую бороздку акцепторного стебля, как показано на РИС. 9.17. Акцепторный стебель при этом сохраняет свою обычную спиральную структуру. АТФ-связывающий участок расположен рядом с концевым аденином. Другим краем своего субстратсвязывающего сайта фермент образует плотный контакт с антикодоновой петлей тРНК, претерпевающей для этого конформационные изменения.

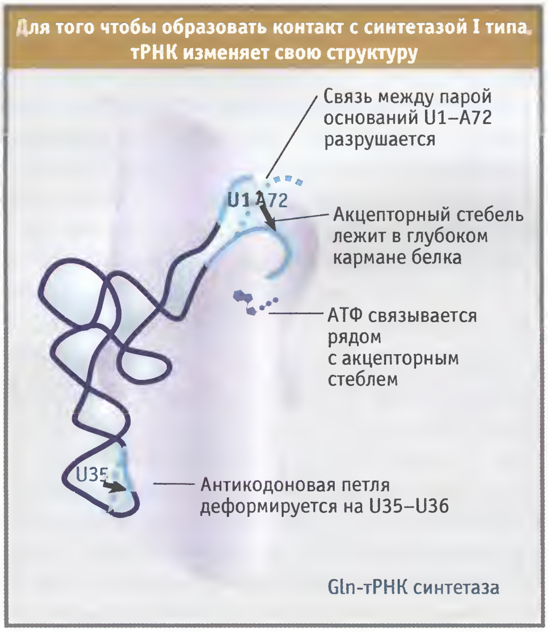


РИС. 9.16 Аминоацил-тРНК-синтетаза I класса контактирует с тРНК в области малой бороздки акцепторного стебля и антикодоновой петли

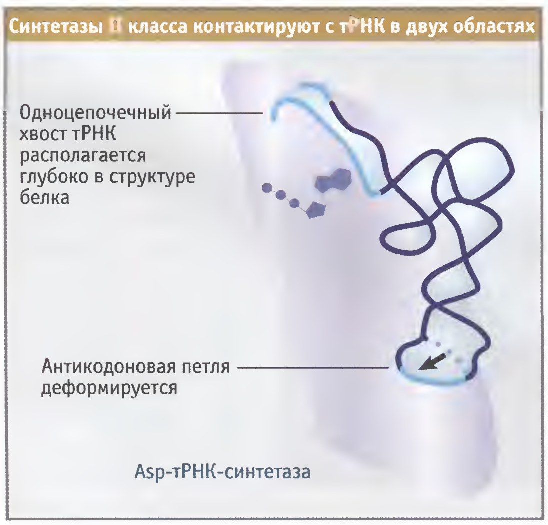


РИС. 9.17 Аминоацил-тРНК-синтетаза II класса контактирует с тРНК в области большой бороздки акцепторного участка и антикодоновой петли

9.11 Синтетазы — ферменты с корректирующей активностью

Ключевые понятия:

- Специфичность реакции узнавания между аминокислотой и тРНК контролируется аминоацил-тРНК-синтетазой, корректирующая активность которой обеспечивает исправление ошибок в случае связывания неправильного компонента.

Аминоацил-тРНК-синтетазы выполняют непростую функцию. Каждая синтетаза должна уметь отличать свою аминокислоту от 20 других, а также распознавать все соответствующие ей тРНК (как правило, от одной до трех) среди общего их многообразия (возможно, общее разнообразие тРНК достигает 100).

Многие аминокислоты похожи друг на друга. Кроме того, большим сходством с аминокислотами обладает ряд промежуточных метаболитов, образующихся в результате химических превращений аминокислот. Особенно сложно провести различия между двумя аминокислотами, отличающимися только длиной углеводородной цепи (например, на одну CH_2-группу). Исходя из относительной энергии связывания двух таких аминокислот, их коэффициент дискриминации должен бы был составлять примерно 1/5. Синтетазы увеличивают этот коэффициент примерно в тысячу раз!

Проводить различия между отдельными тРНК проще, поскольку более крупная структура молекулы тРНК способна формировать большее число точек контакта с ферментом. Однако дело усложняется тем, что все тРНК принимают похожие конформации с ограниченным набором отличительных признаков, позволяющих синтетазе отличать необходимую тРНК от всех прочих.

Существует два основных механизма, описывающих узнавание субстрата ферментом:

- Поступление, проверка правильности, включение или удаление выбранного компонента представляют собой единый акт, предваряющий собственно катализ реакции. Иными словами, в субстратсвязывающий участок может попасть только нужный субстрат. В случае синтетаз это означает, что только правильная аминокислота и соответствующие ей изоакцепторные тРНК обладают способностью стабильно прикрепляться к субстратсвязывающему сайту данного фермента.
- Другой возможный механизм состоит в том, что реакция может пройти ряд стадий, прежде чем будет принято решение о правильности выбранного субстрата. Если субстрат не удовлетворяет нужным критериям, то реакция протекает в обратном (или обходном) направлении, в результате чего неправильный субстрат удаляется. Такой способ проверки (осуществляемый уже после того, как катализируемая реакция прошла) называют **коррекцией**. В случае синтетаз использование такого механизма означает, что реакция соединения аминокислоты и тРНК может катализироваться, даже если в системе присутствуют неподходящие тРНК или аминокислоты.

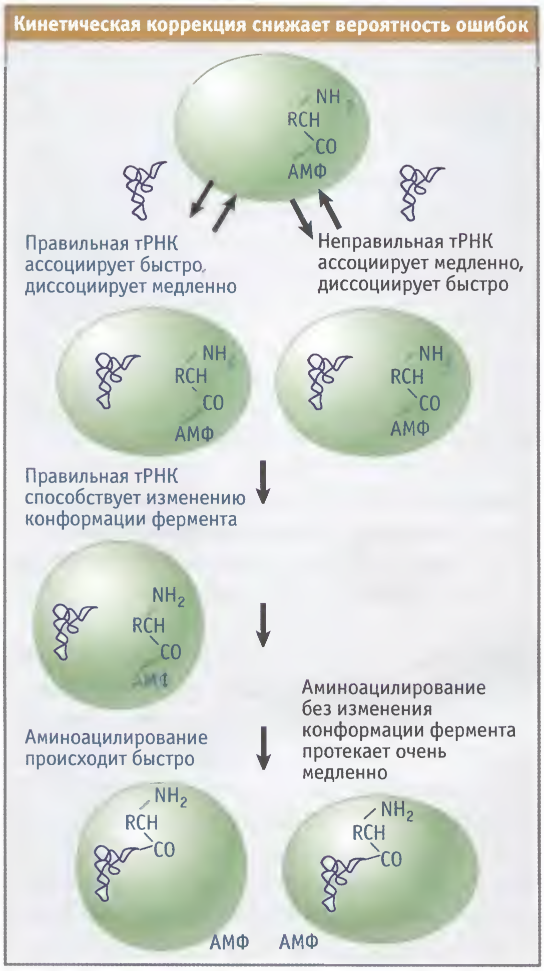


РИС. 9.18 Контроль узнавания синтетазой правильной тРНК включает две стадии. Во-первых, правильная тРНК обладает большей аффинностью к связывающему сайту фермента. Во-вторых, аминоацилирование ошибочной тРНК происходит гораздо медленнее

Синтетазы используют механизмы коррекции при узнавании обоих своих субстратов. С помощью этих механизмов они значительно повышают эффективность выбора как правильных аминокислот, так и правильных тРНК (эффективность дискриминации более или менее постоянна внутри каждой из групп). Больше ошибок синтетазы делают при выборе нужной аминокислоты (величина ошибки колеблется от 10^{-4} до 10^{-5}), при выборе тРНК эта величина составляет около 10^{-6} (см. рис. 8.8). Транспортная РНК взаимодействует с синтетазой по двустадийной схеме, изображенной на РИС. 9.18. Правильная тРНК обладает большей аффинностью к субстратсвязывающему сайту фермента, а потому связывается с ним быстро, а диссоциирует медленно. После савязывания тРНК фермент как бы проводит тщательный анализ. В случае поступления правильной тРНК взаимодействие стабилизируется путем конформационных изменений фермента, в свою очередь разрешающих реакцию быстрого аминоацилирования. Если же поступает неправильная тРНК, конформационных изменений фермента не происходит. В отсутствие конформационных изменений аминоацилирование протекает гораздо медленнее, а следовательно, увеличивается вероятность того, что диссоциация тРНК от фермента произойдет раньше, чем она будет аминоацилирована. Такой тип проверки называют **кинетической коррекцией**.

Способность синтетаз связывать только «свою» аминокислоту различна. Некоторые синтетазы действительно связывают только определенную аминокислоту, тогда как другие способны взаимодействовать с целым рядом похожих по строению аминокислот. Иногда неверно связанные аминокислоты могут даже быть превращены в аденилатную форму, но ни в одном из известных случаев ошибочно активированная аминокислота не образует стабильной аминоацил-тРНК.

Для точного узнавания ферментом аминокислоты необходимо присутствие соответствующей ей тРНК (независимо от того, происходит ли узнавание аминокислоты до стадии формирования аминоацил-аденилата или после). Исключение составляет Met-тРНК-синтетаза, способная отклонить неправильный аминоацил-аденилатный комплекс даже в отсутствие тРНК.

Выделяют два способа коррекции, с помощью которых фермент может удалять неправильную аминокислоту в процессе формирования аминоацил-тРНК. На РИС. 9.19 показано, что прекращение катализа в обоих

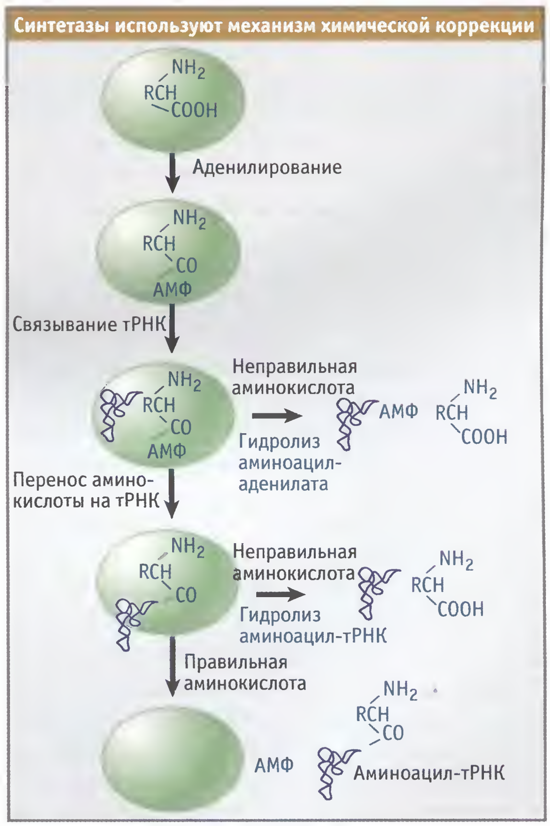


РИС. 9.19 После связывания с ошибочной аминокислотой синтетаза должна связать правильную тРНК. В результате конформационных изменений фермента происходит гидролиз неправильного аминоацил-аденилата. Гидролиз ошибочного комплекса может произойти и после переноса аминокислоты на тРНК

Процесс контролируется на каждой стадии	
Стадия	Частота ошибок
Активация Валина, превращение его в Val-АМФ^{Ile}	1/225
Высвобождение Val–тРНК	1/270
Общая вероятность ошибки	1/225 × 1/270=1/60 000

РИС. 9.20 Точность узнавания тРНК-Ile соответствующим ферментом зависит от контроля возможной ошибки на двух стадиях

случаях является результатом **химической коррекции**. Степень преобладания одного способа над другим у разных ферментов варьирует:

- Неправильный аминоацил-аденилат может гидролизоваться при связывании правильной тРНК. Этот механизм используют несколько ферментов, включая аминоацил-тРНК-синтетазы метионина, изолейцина и валина. (Обычно реакцию нельзя зарегистрировать *in vivo*, однако она продемонстрирована для Met-тРНК-синтетазы в случае неправильной активации аминокислотой гомоцистеином, дефицитным по метильным группам метионина.) В результате коррекции аминокислота высвобождается в измененной форме, такой как тиолактон гомоцистеина. Показано, что тиолактон гомоцистеина образуется в *E.coli* в качестве побочного продукта работы Met-тРНК-синтетазы. Данный факт демонстрирует, что непрерывная коррекция является частью процесса взаимодействия тРНК с соответствующей аминокислотой.
- Некоторые синтетазы используют механизмы химической коррекции на более поздних стадиях. Например, неправильная аминокислота может быть перенесена на тРНК, а затем, находясь уже в тРНК-связывающем сайте фермента, узнается как неправильная по своей структуре, после чего гидролизуется и освобождается. Такой механизм коррекции подразумевает наличие постоянно действующего цикла связывания и гидролиза, продолжающегося до тех пор, пока к тРНК не присоединится правильная аминокислота.

Классическим примером, показывающим, что узнавание аминокислот зависит от присутствия на синтетазе тРНК, является работа Ile-тРНК-синтетазы *E.coli*. Этот фермент способен активировать валин, образуя валиладенилат, но после связывания тРНК^{Ile} фермент гидролизует ранее образовавшийся аденилат. Как показано на РИС. 9.20, итоговая ошибка аминоацилирования зависит от специфичности каждой из описанных стадий. Суммарная ошибка процесса составляет $1{,}5 \times 10^{-5}$, что меньше значений, полученных при измерении частоты замены валина на изолейцин (в глобине кролика), колеблющейся от 2 до 5×10^{-4}. Следовательно, неправильное активирование составляет лишь небольшую долю ошибок, происходящих при синтезе белка в клетке.

Ile-тРНК-синтетаза различает аминокислоты на основе их размеров. Как изображено на РИС. 9.21, фермент имеет два активных центра: активационный (или синтетический) и коррекционный (или гидролитический). Структура кристалла фермента показывает, что синтетический участок слишком мал для того, чтобы обеспечить вхождение туда лейцина (близкого по структуре к изолейцину). Все аминокислоты, размер которых превышает размер изолейцина, не активируются, поскольку не могут проникнуть в синтетический центр фермента. Аминокислоты, которым удалось попасть в синтетический центр, связываются с тРНК, после чего фермент пытается перенести их в коррекционный центр. Изолейцин вследствие своих больших размеров не может войти в коррекционный центр и не подвергается коррекции. Меньший же по размерам валин обладает способностью проникать в коррекционный центр, в результате чего неправильная Val-тРНК^{Ile} гидролизуется. Фактически фермент осуществляет двойное молекулярное распознавание, в ходе которого размер молекулы аминокислоты используется как признак, позволяющий проводить различия между близкими по структуре аминокислотами.

Интересной чертой Ile-тРНК-синтетазы является удаленность синтетического и коррекционного сайтов друг от друга, это расстояние составляет ~34 Å. Структура кристалла фермента в комплексе с подвергнутым коррекции аналогом изолейцина показывает, что аминокислота транспортируется из синтетического сайта в коррекционный. Как показано на РИС. 9.22, этот процесс включает в себя конформационные изменения тРНК. Акцепторный стебель тРНК^{Ile} может существовать в аль-

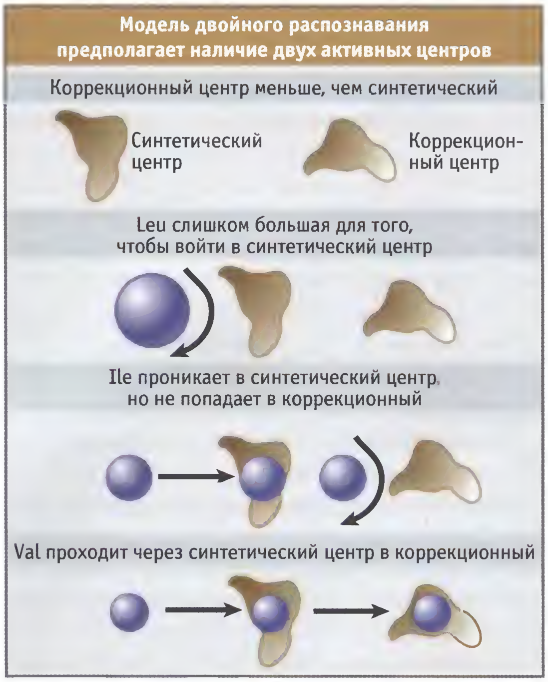


РИС. 9.21 Ile-тРНК синтетаза имеет два активных центра. Аминокислоты, размер которых превышает Ile, не могут быть активированы, поскольку не могут проникнуть в синтетический центр. Аминокислоты, размеры которых меньше размеров Ile, удаляются, так как способны проникать в коррекционный центр

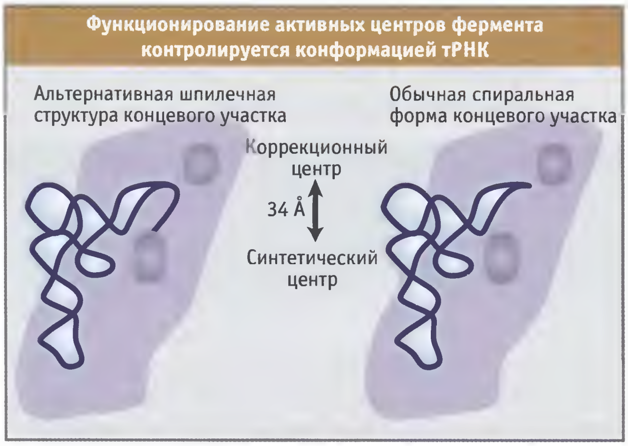


РИС. 9.22 Аминокислота перемещается из синтетического центра Ile-тРНК синтетазы в коррекционный центр в результате изменения конформации акцепторного стебля тРНК

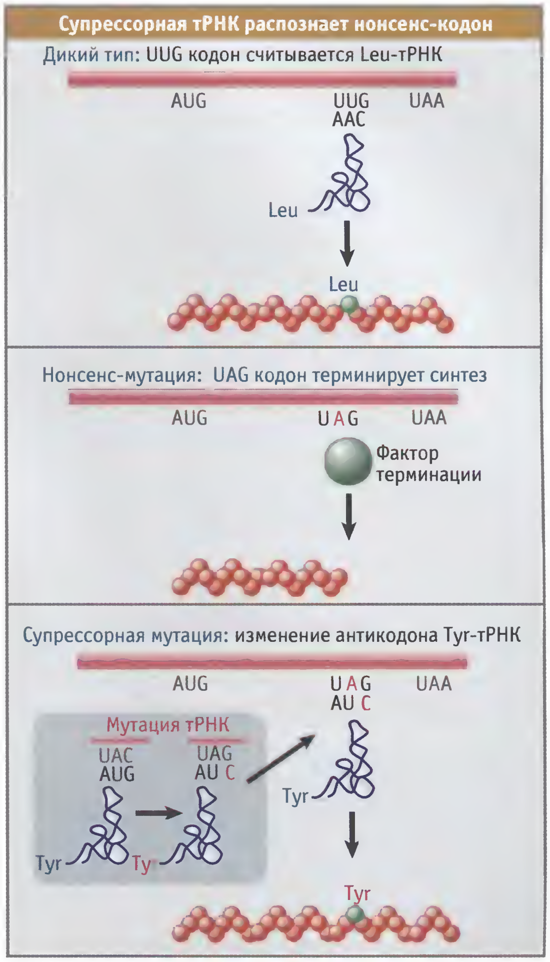


РИС. 9.23 Нонсенс-мутации могут быть супрессированы тРНК с мутантным антикодоном, вставляющей в белок аминокислоту в ответ на мутантный кодон мРНК. В результате синтезируется полноразмерный белок, в котором исходный остаток Leu заменен на Tyr

тернативной конформации. Для того чтобы обеспечить аминоацилирование в синтетическом сайте фермента, он формирует шпильку, после чего возвращается к своей обычной спиральной структуре и перемещает аминокислоту в коррекционный центр. Транслокация между центрами фермента является этапом, лимитирующим скорость коррекции. Ile-тРНК-синтетаза является синтетазой I класса, однако механизм двойного распознавания используется и ферментами II класса.

9.12 Мутантные антикодоны дают возможность супрессорным тРНК считывать новые кодоны

Ключевые понятия:

- Супрессорые тРНК обычно несут мутации в антикодоне, приводящие к изменениям типа кодона, узнаваемого данной тРНК.
- В результате супрессорной мутации в антикодоне тРНК приобретает способность распознавать стоп-кодон как смысловой, обеспечивая включение аминокислоты в полипептидную цепь и синтез белка за стоп-кодон. Это явление получило название нонсенс-супрессии или супрессии преждевременной терминации белкового синтеза.
- Миссенс-супрессия проявляется в способности тРНК узнавать кодоны, отличные от обычных для данной тРНК, и приводит к замене одной аминокислоты на другую.

Мутантные тРНК явились одним из наиболее мощных инструментов при исследовании кодон-антикодонового взаимодействия, а также при выяснении особенностей различных тРНК в этом процессе.

Мутантные тРНК обнаруживают благодаря их способности компенсировать эффект мутаций, присутствующих в генах определенных белков. В терминах классической генетики мутации, способные устранять эффекты, вызванные другими мутациями, называют супрессорными. В системах, на которых изучают супрессорные тРНК, первичная мутация, изменяющая кодон в мРНК, инактивирует белковый продукт. Вторичная (супрессорная) мутация изменяет антикодон тРНК таким образом, что он начинает узнавать мутантный кодон вместо, или так же хорошо, как и свой исходный кодон. В результате включаемые в полипептидную цепь аминокислоты восстанавливают функцию белка. В зависимости от природы первоначальной мутации измененные тРНК получили названия **нонсенс**- или **миссенс-супрессоров**.

В клетках дикого типа нонсен-мутации распознаются только релизинг-факторами, приводящими к терминации синтеза белка. Супрессорные мутации в данном случае придают аминоацил-тРНК новые свойства, позволяющие ей узнавать стоп-кодон как смысловой. Соответственно, включение аминокислоты обеспечивает продолжение синтеза белка за стоп-кодон. Описанный механизм показан на РИС. 9.23. Если аминокислота, включаемая в белок в результате супрессии, отличается от исходно кодируемой ами-

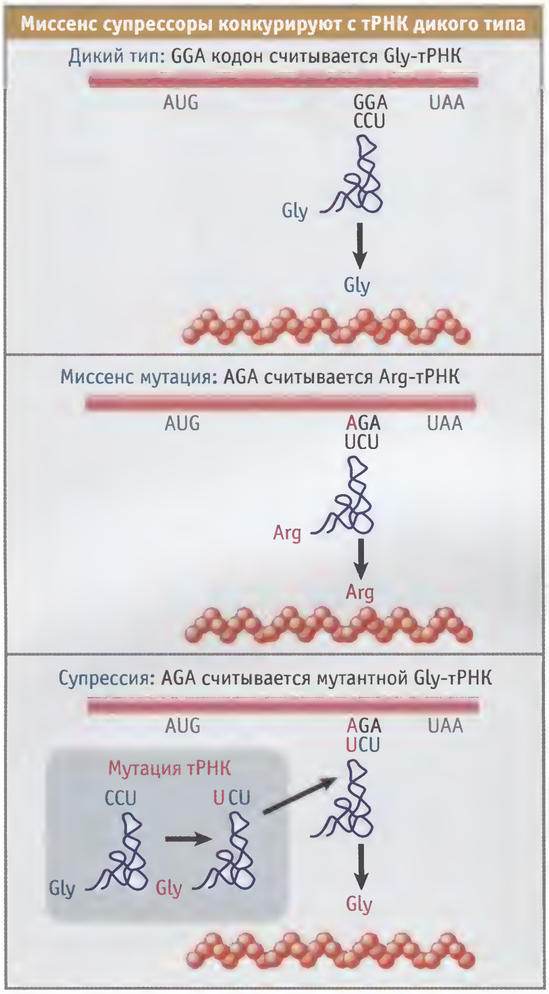


РИС. 9.24 При миссенс-супрессии антикодон тРНК мутирует таким образом, что приобретает способность отвечать на неправильный кодон. Супрессия в этом случае является частичной, поскольку и тРНК дикого типа, и супрессорная тРНК могут взаимодействовать с кодоном AGA

нокислоты в белке дикого типа, то активность белка может измениться.

Миссенс-мутации изменяют смысл кодона, приводя к замене одной аминокислоты на другую, что вызывает образование нефункционального белка (или белка с измененной функцией). (Строго говоря, любая замена аминокислоты в белке является миссенс-мутацией, но на практике мутации обнаруживаются только в том случае, если приводят к изменению активности белка.) Миссенс-мутации супрессируются в результате включения в белок либо исходной, либо похожей на исходную аминокислоты, приводя к восстановлению функции белка. На РИС. 9.24 показано, что миссенс-супрессия осуществляется так же, как и нонсенс-супрессия. В результате мутации в антикодоне тРНК, несущая подходящую аминокислоту, становится способной узнавать мутантный кодон. Таким образом, суть миссенс-супрессии заключается в изменении значения кодона с одной аминокислоты на другую.

9.13 Для каждого терминирующего кодона существуют свои нонсенс-супрессоры

Ключевые понятия:

- Каждому типу нонсенс-кодона соответствует тРНК с определенным мутантным антикодоном.
- Иногда супрессорные мутации возникают не только в антикодонах, но и в других участках молекулы тРНК.

Нонсенс-супрессоры подразделяют на три класса — по одному на каждый стоп-кодон. На РИС. 9.25 описаны свойства некоторых наиболее хорошо изученных супрессоров.

Раньше других были охарактеризованы так называемые амбер-супрессоры. У *E.coli* существует по крайней мере шесть тРНК, которые в результате мутаций начинают узнавать кодон UAG. У всех амбер-супрессорных тРНК антикодоном служит триплет CUA$^{\leftarrow}$. В каждом случае такой триплет возникает в результате единичной замены основания в кодоне дикого типа. Как видно из сравнения последовательностей *supD*, *supE* и *supF*, мутация может затрагивать любое из трех оснований в антикодоне. В итоге, вместо своего прежнего кодона(ов) каждая супрессорная тРНК узнает кодон UAG. При этом в белок включаются серин, глутамин или тирозин — те же аминокислоты, что соответствовали тРНК дикого типа.

Возникновение охра-супрессоров также является результатом мутаций в антикодоне тРНК. Наиболее изученными супрессорами такого типа являются *supC* и *supG*, включающие тирозин или лейцин в ответ на охра (UAA) и амбер (UAG) стоп-кодоны. Такое поведение тРНК полностью согласуется с гипотезой неоднозначного спаривания, по которой кодон UAA не может узнаваться однозначно.

Супрессор кодона UGA обладает неожиданными свойствами. Он возникает в результате мутации в гене

Супрессоры характеризуются наличием мутаций в антикодоне

Локус	тРНК	Дикий тип		Супрессор	
		Кодон/антикодон		Антикодон/кодон	
supD (su1)	Ser	UCG	CGA	CUA	UAG
supE (su2)	Gln	CAG	CUG	CUA	UAG
supF (su3)	Tyr	UA$^{C}_{U}$	GUA	CUA	UAG
supC (su4)	Tyr	UA$^{C}_{U}$	GUA	UUA	UA$^{A}_{G}$
supG (su5)	Lys	AA$^{A}_{G}$	UUU	UUA	UA$^{A}_{G}$
supU (su7)	Trp	UGG	CCA	UCA	UG$^{A}_{G}$

РИС. 9.25 Нонсенс-супрессорные тРНК образуются в результате мутаций в антикодоне

тРНК^{Trp}, заменяющей в 24-м положении A на G. Это приводит к замене пары G-U в D-стебле молекулы на A-U, повышая стабильность спирали. Последовательность антикодона при этом остается той же, что и в тРНК дикого типа: CCA$^{\leftarrow}$. Таким образом, мутация в D-стебле должна каким-то образом изменять конформацию антикодоновой петли, позволяя CCA$^{\leftarrow}$ взаимодействовать с UGA путем образования необычной воблинг-пары C-A. Данная супрессорная тРНК сохраняет способность узнавать и свой обычный кодон UGG.

Аналогичное кодон-антикодоновое взаимодействие обнаруживается и в случае одной из эукариотических тРНК. В печени крупного рогатого скота обнаружена тРНК^{Ser}, антикодоном которой является триплет ^{m}CCA$^{\leftarrow}$. Согласно гипотезе неоднозначного спаривания, такая тРНК должна взаимодействовать с триптофановым кодоном UGG. Но на деле она узнает только стоп-кодон UGA. Таким образом, можно сказать, что UGA в описанной ситуации как бы «естественно супрессирован».

Важность приведенных наблюдений состоит в том, что они продемонстрировали невозможность исчерпывающего описания кодон-антикодонового взаимодействия только на основе триплетной последовательности (причем как для тРНК дикого типа, так и для мутантных тРНК). Было доказано влияние на процесс узнавания и других особенностей молекулы тРНК.

9.14 Супрессорные тРНК могут конкурировать с тРНК дикого типа за связывание с кодонами

Ключевые понятия:

- Супрессорные тРНК и тРНК дикого типа, обладающие одинаковыми антикодонами, конкурируют друг с другом за возможность взаимодействия с соответствующим кодоном.
- Негативный эффект супрессии заключается в считывании и правильных стоп-кодонов.
- Кодон UGA даже в отсутствие мутантной супрессорной тРНК может узнаваться Trp-тРНК с частотой 1–3%.

Существует любопытное различие между ситуациями, когда кодон узнается своей собственной аминоацил-тРНК, и когда мутация позволяет супрессорной тРНК узнавать новый кодон. В клетке дикого типа каждому кодону может быть присвоено только одно значение, соответствующее либо конкретной аминокислоте, либо сигналу терминации трансляции. В клетке же, несущей супрессорную мутацию, мутантный кодон имеет два значения: он может быть узнан супрессорной тРНК либо быть прочитан согласно своему обычному значению.

Нонсенс-супрессорная тРНК должна, таким образом, конкурировать с релизинг-факторами за терминирующий кодон. А миссенс-супрессорная тРНК должна конкурировать с «оригинальной» тРНК. Степень конкуренции влияет на эффективность супрессии, приводя к зависимости этого процесса не только от степени сродства между кодоном и антикодоном, но и от концентрации супрессоров в клетке, а также от других параметров, влияющих на протекание этих конкурирующих реакций.

Так, на эффективность считывания кодона оказывает влияние его окружение, и степень нонсенс-супрессии молекулами определенной тРНК колеблется в широких пределах в зависимости от соседствующих оснований. Мы не знаем механизмов, посредством которых окружение кодона оказывает влияние на кодон-антикодоновое взаимодействие. Но в зависимости от локализации, частота узнавания кодона определенной тРНК может изменяться больше, чем на порядок. Наибольший эффект при этом оказывает основание, примыкающее к 3′-концу кодона.

Нонсенс-супрессоры были выделены благодаря их способности узнавать мутантные нонсенс-кодоны. Однако в клетке есть и нормальный стоп-кодон с такой же последовательностью. Мутантная тРНК, супрессирующая нонсенс-мутацию в принципе должна супрессировать и соответствующий нормальный стоп-кодон в конце каждого гена, заканчивающегося на этот кодон. На РИС. 9.26 показано, как в результате такого **сквозного прочитывания** образуется более

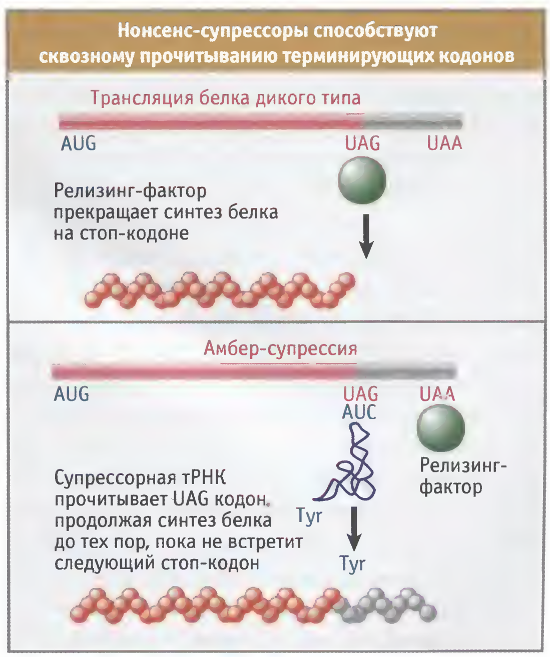


РИС. 9.26 Нонсенс-супрессоры обладают способностью прочитывать нормальные терминирующие кодоны. При этом образуется более длинный белок, чем белок дикого типа

длинный белок с дополнительной С-концевой последовательностью. Синтез таких протяженных молекул белка будет заканчиваться на следующем (отличном от супрессированного) стоп-кодоне в этой же рамке считывания. Любая эффективная супрессия терминации негативно скажется на жизнедеятельности клетки, так как при этом будут образовываться более длинные белки с измененной функцией.

Эффективность роста клеток с амбер-супрессией весьма высока и составляет, в зависимости от конкретной системы, от 10 до 50% (от нормы). Столь высокая эффективность роста возможна благодаря тому, что амбер-кодоны используются для терминации белкового синтеза у *E.coli* довольно редко.

Обнаружить охра-супрессоры сложнее. Они всегда характеризуются более низкой эффективностью роста — обычно ниже 10%. Это является следствием того, что одновременная супрессия кодонов UAA и UAG неблагоприятна для клеток *E.coli,* возможно, в силу того, что охра-кодоны используются в качестве природных сигналов терминации наиболее часто.

Кодон UGA по своей природе является наименее эффективным терминирующим кодоном; Trp-тРНК способна ошибочно считывать его в клетках дикого типа с частотой 1–3%. Однако, несмотря на это, он используется для терминации у бактерий чаще, чем амбер-триплеты.

Миссенс-супрессор одного гена, скорее всего, будет вызывать мутации и в других генах. Если супрессор исправляет эффект мутации, заменяя одну аминокислоту на другую в мутантном сайте, он будет вводить новую аминокислоту взамен обычной и в других аналогичных сайтах, а такие изменения могут подавлять нормальную функцию белка.

Поэтому перед клеткой возникает дилемма: с одной стороны, она должна супрессировать мутантный кодон, с другой — не допустить слишком сильных изменений нормальных значений кодонов в других местах. Отсутствие сколько-нибудь сильных миссенс-супрессоров объясняется негативным воздействием, которое они оказывают на другие гены.

Мутация, вызывающая образование супрессорной тРНК, может иметь два следствия. Во-первых, она дает возможность тРНК узнавать новый кодон. Во-вторых, в некоторых случаях она не позволяет тРНК взаимодействовать со своим прежним кодоном. Важно отметить, что все эффективные амбер-супрессоры возникают в результате мутирования одной из «продублированных» тРНК. В этом случае клетка попрежнему будет иметь тРНК, взаимодействующие с исходным кодоном. Следовательно, мутация не выключит узнавание прежнего кодона, который будет обслуживаться оставшимися синонимичными тРНК. В редких случаях, когда с кодоном взаимодействует единственная молекула тРНК, любая мутация, нарушающая это взаимодействие, является летальной.

Чаще всего супрессия рассматривается в контексте возникновения мутаций, изменяющих смысл кодона. Однако с небольшой частотой стоп-кодоны прочитываются как смысловые и в клетках дикого типа. Первый такой пример был обнаружен в гене поверхностного белка фага Qβ. Слабая супрессия стоп-кодона, расположенного на конце этого гена, необходима для того, чтобы при формировании инфекционных частиц фага Qβ продуцировалось некоторое количество поверхностного белка с дополнительной С-концевой последовательностью. Действительно, этот стоп-кодон является слабым (подтекающим, leaky), поскольку с определенной вероятностью узнается Trp-тРНК.

Сквозное прочитывание стоп-кодонов встречается и у эукариот, где это явление часто используется вирусами с РНК геномом. Данный процесс включает в себя супрессию UAG/UAA молекулами Tyr-тРНК, Gln-тРНК и Leu-тРНК, или же супрессию UGA молекулами Trp-тРНК или Arg-тРНК. Степень частичной супресии определяется контекстом окружения кодона.

9.15 Рибосомы влияют на точность трансляции

Ключевые понятия:

- Структура 16s-рРНК в рибосомальных сайтах А и Р влияет на точность трансляции.

При анализе первичной структуры белка не удается обнаружить каких-либо вариаций в последовательности аминокислот, что свидетельствует о высочайшей точности белкового синтеза. Очень небольшое число ошибок проявляется в виде замены одной аминокислоты на другую. Эти ошибки могут возникать на двух стадиях синтеза белка (см. рис. 8.8 в разд. 8.3 «Специальные механизмы контролируют точность синтеза белка»):

- Критичным моментом является «зарядка» тРНК единственно правильной аминокислотой. Это — функция аминоацил-тРНК-синтетазы. Частота ошибки на данном этапе различается для разных ферментов, однако в среднем это 1 на 10^5 случаев аминоацилирования.
- Специфичность кодон-антикодонового взаимодействия является вторым решающим фактором. Хотя для разных кодон-антикодоновых пар значения констант связывания могут различаться, специфичность всегда слишком низка, чтобы обеспечить уровень ошибок менее 10^{-5}. В растворе свободная тРНК довольно слабо связывается с соответствующим ей триплетом. Похожие, но не совпадающие триплеты (имеющие два правильных основания из трех) узнаются лишь в 10–100 раз менее эффективно, чем правильные триплеты.

Таким образом, кодон-антикодоновое узнавание является как бы слабым звеном в процессе обеспече-

ния точности трансляции. Однако рибосомы, играющие важнейшую роль в контроле за специфичностью этого взаимодействия, отличают корректные кодон-антикодоновые пары от некорректных и повышают эффективность отбора нужных пар примерно в 1000 раз. В дополнение к корректирующей функции самих рибосом, на точность узнавания кодона антикодоном оказывают влияние белковые факторы, доставляющие инициаторную и все остальные аминоацил-тРНК на рибосомы.

Следовательно, должен существовать некий механизм, ответственный за стабилизацию правильной аминоацил-тРНК на рибосоме и позволяющий включить доставленную аминокислоту в синтезируемую полипептидную цепь; в то же время контакт рибосомы с некорректной аминоацил-тРНК должен быстро разрушаться (до того, как комплекс успеет прореагировать). Предположим, что взаимодействие тРНК с рибосомой (в составе аминоацил-тРНК–EF-Tu–ГТФ комплекса) не является специфичным. При условии, что любой комплекс (независимо от природы входящей в его состав тРНК) может попасть в А-сайт, число неправильных взаимодействий должно значительно превышать число правильных.

Существуют две основные модели, согласно которым рибосомы дискриминируют верные и ошибочно привнесенные аминоацил-тРНК. Вероятнее всего, на практике происходит комбинация элементов обеих моделей.

- Модель прямого распознавания предполагает, что сама структура рибосомы приспособлена для распознавания точно спаренных аминоцил-тРНК. Иными словами, в результате корректного кодон-антикодонового взаимодействия возникают незначительные изменения конформации молекулы аминоацил-тРНК, которые и распознает рибосома. В этой модели процесс распознавания правильной тРНК предшествует всем остальным реакциям.
- Модель кинетической коррекции предполагает по крайней мере две стадии, на которых может происходить селекция аминоацил-тРНК. Следовательно, неправильная аминоацил-тРНК может пройти ряд этапов прежде, чем будет отклонена. Общая эффективность отбора нужных аминоацил-тРНК является в этом случае результатом селекции на каждой стадии процесса.

На **РИС. 9.27** схематично изображено, что происходит с правильными и неправильными аминоацил-тРНК на рибосоме. Правильная аминоацил-тРНК образует стабильные контакты с рРНК. Ошибочная аминоацил-тРНК не способна формировать такие контакты и поэтому легко диффундирует из А-сайта рибосомы.

Исследование взаимодействия тРНК с рибосомой началось в 1960 году с изучения эффекта антибиотика стрептомицина. Стрептомицин ингибирует синтез белка, связываясь с 16S рРНК и подавляя способность EF-G катализировать транслокацию. К тому же он увеличивает уровень ошибочного считывания пиримидинов U и C (обычно один пиримидин заменяется на другой, изредка — на A). Мишенью, чувствительной к действию стрептомицина, оказался рибосомальный белок S12 (мутация, повышающая устойчивость клеток к действию стрептомицина, влияла на аминокислотную последовательность белка S12). Рибосомы с мутантным белком S12, выделенные из устойчивых к действию антибиотика бактерий, демонстрировали снижение уровня ошибок трансляции по сравнению с рибосомами «дикого типа». Оказалось, что S12 принимает участие в контроле над уровнем ошибок трансляции. Его мутантная форма уменьшая процент ошибок, снижает и эффект стрептомицина.

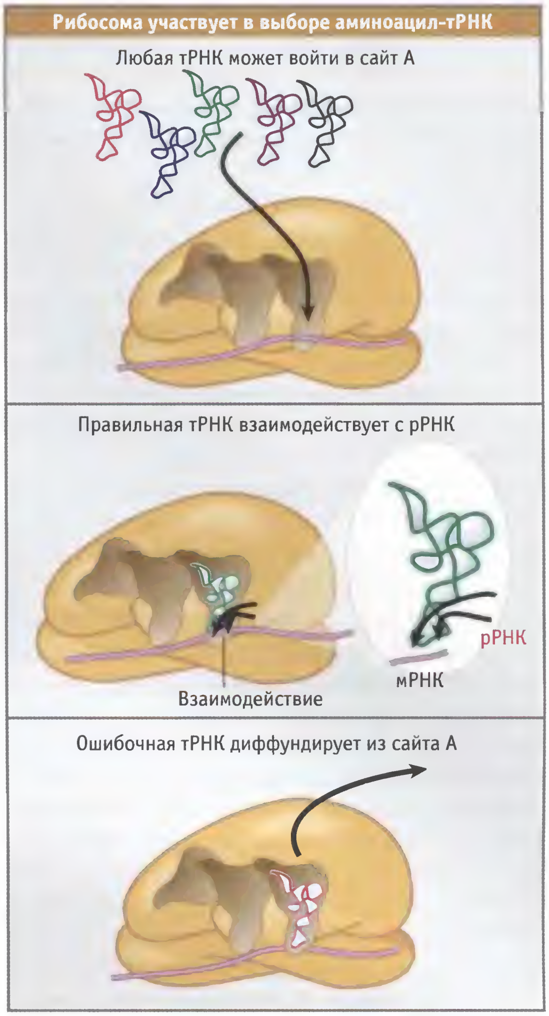


РИС. 9.27 Любая аминоацил-тРНК (с помощью EF-Tu) может разместиться в А-сайте рибосомы, но только та аминоацил-тРНК, антикодон которой провзаимодействует с кодоном мРНК, образует стабильные контакты с рРНК. В случае если эти контакты не образуются, аминоацил-тРНК диффундирует из сайта А рибосомы

Рибосомальный белок S12 стабилизирует структуру 16S рРНК в участке, с которым связывается стрептомицин. *Необходимо отметить, что на точность трансляции влияет P/A-регион рибосомы: точность транс-*

ляции меняется с изменением структуры 16S рРНК. Комбинация эффектов белка S12 и стрептомицина на структуру рРНК объясняет фенотип различных мутантов по S12. У некоторых из них возможность корректной работы рибосом даже становится зависимой от наличия стрептомицина.

Сейчас, исходя из структуры кристалла рибосомы, мы знаем, что 16S рРНК отвечает за формирование контактов с аминоацил-тРНК. Два основания 16S рРНК могут контактировать с малой бороздкой спирали, формируемой антикодоном тРНК и первыми двумя основаниями в кодоне мРНК. Это оказывает прямой стабилизирующий эффект на взаимодействие кодона и антикодона в первых двух позициях кодона, однако не затрагивает третью позицию.

Стабилизация корректно спаренной аминоацил-тРНК обладает двумя следствиями: во-первых, удержание аминоацил-тРНК в А-сайте рибосомы не позволяет ей уклониться от последующих стадий белкового синтеза; во-вторых, конформационные изменения рРНК запускают следующую стадию — гидролиз ГТФ с помощью фактора EF-Tu.

Точность трансляции связана со скоростью работы рибосомы. Пока аминоацил-тРНК находится в сайте А рибосомы, она может быть вовлечена в последующие стадии белкового синтеза. Таким образом, пауза между попаданием аминоацил-тРНК в А-сайт и последующей реакцией транспептидации дает дополнительную возможность для диссоциации ошибочной аминоацил-тРНК. Диссоциация ошибочной аминоацил-тРНК происходит гораздо быстрее диссоциации подходящей аминоацил-тРНК (приблизительно в 5 раз). Поэтому чем ниже скорость транспептидации, тем выше вероятность диссоциации неправильной аминоацил-тРНК.

Ранее контроль за специфичностью узнавания аминоацил-тРНК приписывали самой рибосоме, однако результаты последних работ позволяют считать, что на этот процесс и в сайтах А и Р существенное влияние оказывают факторы транслокации. Существуют данные, показывающие, что фактор EF-Tu вовлечен в поддержание правильной рамки считывания. Это продемонстрировано на мутантных формах EF-Tu, супрессирущих сдвиг рамки. Предполагается, что EF-Tu не только доставляет аминоацил-тРНК в сайт А рибосомы, но и участвует в ориентации только что поступившей аминоацил-тРНК относительно пептидил-тРНК в сайте Р.

Наиболее ярким примером влияния белковых факторов трансляции является их способность изменять значение кодона при инициации синтеза белка. У дрожжей в гене HIS4 обнаружены мутации инициирующего кодона AUG, изменяющие его на UUG и тем самым нарушающие инициацию. Однако существуют внешние супрессорные мутации, позволяющие кодону UUG выступать в роли инициирующего. Две таких мутации выявлены в генах, кодирующих α- и β-субъединицы eIF2 — фактора, связывающего Met-тРНК$_i$ с сайтом Р. Мутация в eIFβ2 расположена в участке белка, ответственном за связывание с нуклеиновой кислотой. Предполагается, что мишенью связывания для этого белка является инициирующая последовательность мРНК или же пары оснований, образующиеся между кодоном мРНК и антикодоном Met-тРНК$_i$. Вероятно, помимо доставки инициаторной тРНК в Р-сайт рибосомы, eIF2 участвует и в распознавании инициирующего кодона.

Энергетические затраты на синтез белка увеличиваются за счет дополнительных расходов энергии на процессы коррекции. Вопрос, имеющий важное значение при подсчете энергетических затрат на синтез белка, заключается в следующем: на какой стадии процесса принимается решение о возможности акцептирования тРНК. Если решение принимается сразу после присоединения комплекса аминоацил-тРНК–EF-Tu–ГТФ, то дополнительные затраты энергии (требующиеся для удаления всех ошибочных аминоацил-тРНК, успевающих попасть в сайт А до того, как туда попадет правильная аминоацил-тРНК) будут минимальны. Но если такое решение принимается после гидролиза ГТФ, расходы энергии будут значительно выше. Ошибочная аминоацил-тРНК может диссоциировать до или после гидролиза ГТФ, но когда именно это происходит, до сих пор точно не установлено. Существует ряд фактов, свидетельствующих, что расход ГТФ *in vivo* превышает три макроэргические связи, необходимые для включения каждой (корректной) аминокислоты в полипептидную цепь.

9.16 Перекодирование меняет значение кодона

Ключевые понятия:

- Изменение смысла кодона может быть следствием возникновения мутантных тРНК или тРНК со специфичными свойствами.
- В результате сдвига или пропуска фрагмента рибосома может менять рамку считывания, оба этих процесса зависят от структуры и свойств мРНК.

Как правило, у зрелых мРНК существует единственная непрерывная рамка считывания. Трансляция начинается с кодона AUG и продолжается до первого стоп-кодона в этой рамке. При этом процесс считывания триплетов идет подряд: вставка или делеция одного основания вызывают мутацию (сдвиг рамки считывания). В результате рамка считывания за местом вставки или делеции меняется. Рибосома и тРНК, продолжая считывать триплеты, синтезируют при этом совсем другую последовательность аминокислот.

Существуют ситуации, когда даже в случае сдвига рамки считывания или включения в последовательность нонсенс-кодона, синтез полноразмерного белка не нарушается. Исключения из обычных правил транс-



РИС. 9.28 В результате мутации в индивидуальной тРНК (обычно в антикодоне) процесс узнавания соответствующего ей кодона может супрессироваться. В особых случаях особая тРНК, связанная с необычным фактором элонгации, приобретает способность распознавать терминирующий кодон, прилегающий к шпилечной структуре мРНК

ляции возможны при **перекодировании**. Эти исключения могут быть нескольких типов.

Изменение смысла одного-единственного кодона может вызвать замену одной аминокислоты на другую или включение аминокислоты в ответ на терминирующий кодон. Как показано на РИС. 9.28, такие изменения основываются на особенностях индивидуальных тРНК, узнающих данный кодон:

- Супрессия заключается в распознавании кодона мутантной тРНК, которая в обычных условиях взаимодействует с другим кодоном (см. разд. 9.12 «Мутантные антикодоны дают возможность супрессорным тРНК считывать новые кодоны»).
- Изменение смысла кодона происходит при модификации аминоацил-тРНК. (см. разд. 9.8 «Некоторые стоп-кодоны могут встраивать в белок нестандартные аминокислоты»).

Изменение же рамки считывания также возможно в двух ситуациях:

- Как правило, сдвиг рамки считывания происходит в результате проскальзывания аминоацил-тРНК на одно основание, +1 вперед или -1 назад (см. следующий раздел «Сдвиг рамок считывания происходит на скользких последовательностях»). В результате, как изображено на РИС. 9.29, трансляция может продолжаться за пределы стоп-кодона.
- Двигаясь вдоль мРНК, рибосома способна «перескакивать» с одного кодона на другой, идентичный первому, даже если они разделены протяженной последовательностью. Нуклеотидная последовательность, пропущенная рибосомой, не транслируется в аминокислотную. Как показано на РИС. 9.30, данный процесс позволяет трансляции «перескочить» через любые терминирующие кодоны, расположенные в пропущенной последовательности.

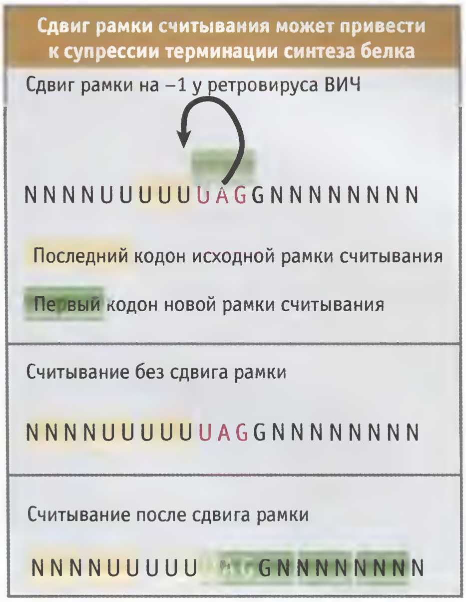


РИС. 9.29 Проскальзывание тРНК на одно основание при связывании с кодоном вызывает сдвиг рамки считывания, что, в свою очередь, приводит к супрессии терминации. Эффективность этого процесса, как правило, составляет ~ 5%

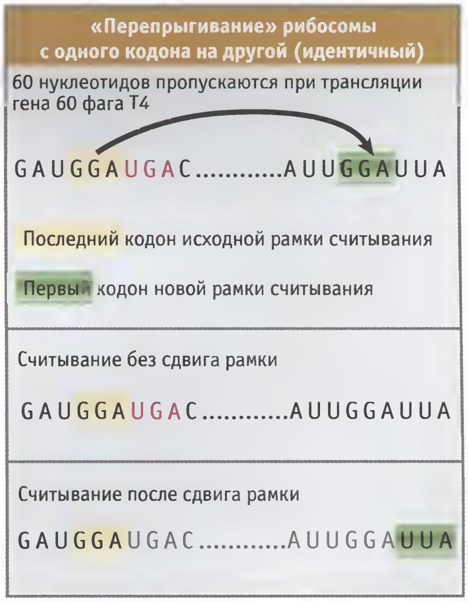


РИС. 9.30 Двигаясь вдоль мРНК, рибосома может пропустить значительный фрагмент, перескочив с одного кодона на другой, в этом случае пептидил-тРНК в сайте Р провзаимодействует со следующим аналогичным кодоном последовательности

9.17 Сдвиг рамок считывания происходит на «скользких» последовательностях

Ключевые понятия:

- Рамка считывания определяется последовательностью нуклеотидов в мРНК и рибосомальным окружением.
- «Скользкие» последовательности способствуют сдвигу тРНК, спаренной с антикодоном, на одно основание, изменяя тем самым рамку считывания.
- Трансляция некоторых генов зависит от регулярных «запрограммированных» сдвигов рамки считывания.

Сдвиг рамки считывания определяется свойствами отдельных тРНК:

- Некоторые мутантные супрессорные тРНК способны узнавать квадриплетный «кодон» (вместо обычного триплета).
- «Скользкие» последовательности способствуют сдвигу тРНК в сайте А на одно основание назад или вперед вдоль последовательности мРНК.

Мутанты со сдвигом рамки считывания образуются в результате вставки или делеции одного основания. Такие мутации могут быть супрессированы за счет восстановления исходной рамки считывания. Эффект может быть достигнут путем компенсаторных вставок или делеций внутри данного гена (см. разд. 2.8 «Генетический код триплетен»). Кроме того, в роли внешних супрессоров мутаций со сдвигом рамки считывания могут выступать тРНК с аберрантными свойствами.

Самым простым примером внешней супрессии мутаций со сдвигом рамки является случай, когда супрессор исправляет сдвиг, вызванный вставкой дополнительного основания в участок, состоящий из нескольких идентичных остатков. Например, основание G может быть включено в участок из нескольких расположенных подряд G. Супрессором такой мутации служит тРНК^{Gly}, содержащая дополнительное основание в антикодоновой петле. В результате антикодон превращается из обычной триплетной последовательности CCC$^{\leftarrow}$ в квадриплетную последовательность CCCC$^{\leftarrow}$. Таким образом супрессорная тРНК узнает кодон, состоящий из четырех оснований.

Некоторые тРНК, супрессирующие мутации со сдвигом рамки, могут взаимодействовать более чем с одним квадриплетным «кодоном». Например, бактериальная супрессорная тРНК^{Lys} способна узнавать кодоны AAAA и AAAU вместо обычного кодона AAA. Другой супрессор может прочитывать любой квадриплетный «кодон», содержащий в трех первых положениях ACC; природа добавочного основания не влияет на процесс узнавания. Во всех этих случаях альтернативные основания, находящиеся в четвертом положении удлиненного «кодона», не имеют отношения к гипотезе неоднозначного спаривания. Вероятно, в действительности такая супрессорная тРНК, как все тРНК, узнает триплетный кодон, но по какой-то причине (возможно, из-за пространственных ограничений) соседнее основание оказывается блокированным. В результате следующая тРНК может провзаимодействовать с кодоном, только пропустив одно основание.

Случаи, когда сдвиг рамки присутствует в норме, характерны для фагов и вирусов. Это явление оказывает влияние на терминацию синтеза белка и являться результатом особенностей структуры вирусных мРНК.

У ретровирусов трансляция первого гена, как и у всех, терминируется стоп-кодоном. Второй ген вируса (кодирующий как бы дополнительную С-концевую последовательность первого белка) лежит сразу за первым геном в другой рамке считывания и (у некоторых вирусов) транслируется в результате сдвига исходной рамки, приводящего к «исчезновению» стоп-кодона (см. рис. 9.29 и разд. 22.3 «Ретровирусные гены, кодирующие полипротеины»). Эффективность сдвига рамки невелика, и, как правило, не превышает 5%. В действительности этот механизм играет важную роль в биологии вирусов, а возрастание эффективности может оказаться даже губительным. На **РИС. 9.31** проиллюстрирована похожая ситуация, характерная для Ty элемента дрожжей. В этом случае для трансляции гена *tyb*, расположенного за геном *tya*, в результате сдвига рамки считывания должен быть пропущен стоп-кодон гена *tya*.

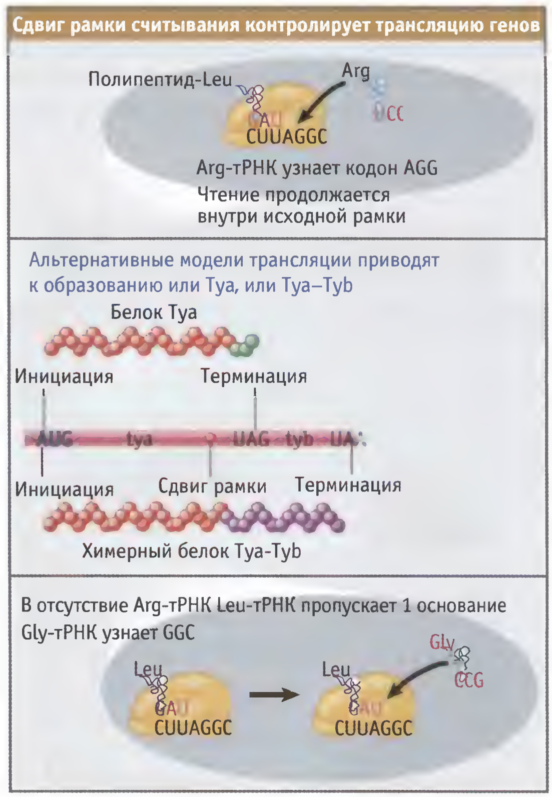


РИС. 9.31 Экспрессия гена *tyb* Ty элемента дрожжей возможна только в случае сдвига рамки считывания на +1. Сдвиг происходит на последовательности из семи оснований, представленной кодоном(ами) Leu, за которым следует кодон редко встречающейся Arg-тРНК

Эти примеры демонстрируют, что редкие (но запланированные) ошибки считывания на самом деле могут являться необходимым компонентом естественного процесса трансляции. Данное явление получило название **запрограммированного сдвига рамки считывания.** Оно происходит в определенных сайтах с частотой, в 100–1000 раз превышающей обычную частоту ошибки (~3×10^{-5} на кодон).

Механизм запрограммированного сдвига рамки характеризуется двумя общими чертами:

- «Скользкие» последовательности позволяют аминоацил-тРНК взаимодействовать со своим кодоном, а затем, сдвигаясь на одно основание вперед (сравнительно редко) либо на одно основание назад (чаще), образовывать пары с еще одним перекрывающимся триплетом, который также может быть спарен со своим антикодоном.
- Рибосомы задерживаются в сайте, где должен произойти сдвиг рамки, для того чтобы дать аминоацил-тРНК возможность перестроить кодон-антикодоновую пару. Причиной такой задержки может быть смежный кодон, требующий редкой аминоацил-тРНК, терминирующий кодон, медленно узнающийся своим белковым фактором освобождения, или препятствия, создаваемые пространственной структурой мРНК (например, псевдоузел (pseudoknot) — особая конформация молекулы РНК), которая задерживает рибосому.

Проскальзывание может осуществляться в обоих направлениях; сдвиг на –1 происходит в том случае, когда тРНК сдвигается назад (к 5′-концу), а сдвиг +1 — когда тРНК движется вперед (к 3′-концу). В обоих случаях результатом является позиционирование следующей аминоацил-тРНК в сайте А вне фазы. Сдвиг рамки осуществляется до момента образования пептидной связи. Чаще всего, если следом за «скользкой» последовательностью, запускающей процесс сдвига рамки, следует вторичная (формирующая шпильку) структура мРНК, окружающие последовательности оказывают влияние на эффективность процесса.

На рис. 9.31 изображена схема участия типичной «скользкой» последовательности в механизмах сдвига рамки считывания. Последовательность из семи нуклеотидов CUUAGGC распознается следующим образом: Leu-тРНК узнает триплет CUU, затем Arg-тРНК узнает триплет AGG. Однако Arg-тРНК встречается редко, и если этот факт приводит к задержке синтеза, Leu-тРНК соскальзывает с CUU кодона на перекрывающийся триплет UUA. Это вызывает сдвиг рамки считывания, поскольку следующий триплет в фазе (GGC) прочитывает Gly-тРНК. Проскальзывание обычно происходит в сайте Р (когда Leu-тРНК превращается в пептидил-тРНК, несущую синтезируемую полипептидную цепь).

Сдвиг рамки на стоп-кодоне вызывает дальнейшее считывание белка. Основание с 3′-конца стоп-кодона влияет на относительную вероятность терминации или сдвига рамки, определяя, таким образом, эффективность сигнала терминации. Этот факт помогает объяснить значение контекста для процесса терминации.

9.18 «Перепрыгивание» через основания осуществляется перемещением рибосомы

Определенные последовательности индуцируют сдвиг рибосомы, в ходе которого рибосома останавливает трансляцию и скользит вдоль мРНК, удерживая пептидил-тРНК в сайте Р. Затем трансляция возобновляется (см. рис. 9.30). Этот феномен довольно редок; обнаружено всего три примера его использования в живых системах, наиболее ярким из которых является «проскок» в гене 60 фага Т4. Осуществляя трансляцию, рибосома сдвигается по мРНК на 60 нуклеотидов.

Ключом к пониманию механизма таких сдвигов являются идентичные (или синонимичные) кодоны на обоих концах пропущенной последовательности. Их иногда называют точками «взлета» и «посадки». До того, как осуществить сдвиг, рибосома занимает позицию, в которой пептидил-тРНК спарена с кодоном «взлета» в сайте Р, а сайт А, оставаясь пустым, вакантен для входа аминоацил-тРНК. На **РИС. 9.32** показано, как рибосома в таком состоянии скользит вдоль мРНК до тех пор, пока пептидил-тРНК не провзаимодействует с

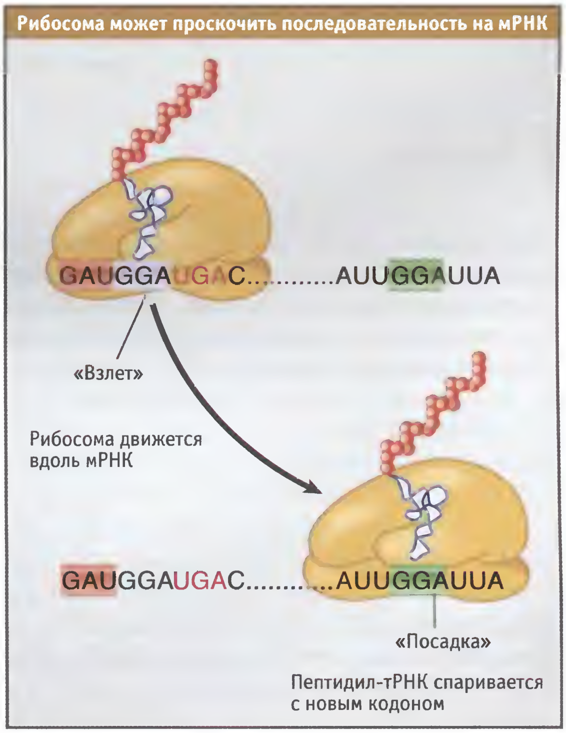


РИС. 9.32 Рибосома, в сайте Р которой находится пептидил-тРНК (спаренная с кодоном мРНК), может приостановить трансляцию. Она скользит вдоль мРНК до тех пор, пока не встретит синонимичный кодон, с которым пептидил-тРНК, все еще находящаяся в сайте Р, сможет провзаимодействовать. После этого синтез белка возобновляется

кодоном в сайте «посадки». Следует отметить высокую эффективность работы такой системы, составляющую около 50%.

Сама последовательность мРНК индуцирует сдвиг рибосомы. Основными параметрами системы, изображенной на рис. 9.32, являются два GGA кодона (для «взлета» и «посадки»), пространство между ними, структура стебель-петля, в состав которой входит кодон «взлета», и примыкающий к кодону «взлета» стоп-кодон. Кроме того, в состав этой системы включен и синтезируемый в данный момент белок.

Стадия «взлета» заключается в разрушении связи между пептидил-тРНК и спаренным с ней кодоном. Последующее движение мРНК предотвращает их повторное взаимодействие. Рибосома сканирует мРНК до тех пор, пока пептидил-тРНК не спарится с кодоном «посадки». Возобновление синтеза белка происходит после того, как новая аминоацил-тРНК войдет в сайт А (по обычному механизму).

Как и сдвиг рамки считывания, сдвиг рибосомы зависит от скорости работы рибосомы. Вероятность диссоциации пептидил-тРНК из сайта Р увеличивается в результате задержки появления аминоацил-тРНК в А-сайте. Нехватка какой-либо аминоацил-тРНК может запустить реакцию сдвига рибосомы в бактериальных системах вследствие длительной паузы на этапе связывания аминоацил-тРНК с сайтом А. В гене 60 фага Т4 особенности структуры мРНК влияют на эффективность терминации, уменьшая ее и тем самым создавая паузу, необходимую для осуществления реакции сдвига.

9.19 Резюме

Последовательность мРНК, прочитываемая триплетами по направлению от 5′- к 3′- концу, посредством генетического кода переводится в аминокислотную последовательность белка, синтезируемого от N- к C-концу. Из 64 возможных триплетов 61 кодируют аминокислоты, а 3 являются сигналами терминации трансляции. Синонимичные кодоны, кодирующие одну и ту же аминокислоту, часто отличаются только основанием в третьем положении триплета. Вырожденность третьего основания, а также тенденция к сходству кодонов, кодирующих схожие по свойствам аминокислоты, снижают эффект мутаций. Генетический код универсален и свой современный вид приобрел уже на очень ранних стадиях эволюции. Отклонения генетического кода от стандарта в ядерном геноме очень редки, однако характерны для митохондриального генома.

С одним кодоном может взаимодействовать несколько разных тРНК. Для каждого организма характерен свой набор тРНК, взаимодействующих с группой синонимичных кодонов. Кодон-антикодоновое взаимодействие характеризуется неоднозначным соответствием в первом положении кодона (третьем положении антикодона), что позволяет некоторым тРНК узнавать несколько кодонов. Все тРНК содержат модифицированные основания, по некоторым из которых аминоацил-тРНК-синтетазы идентифицируют тРНК. На кодон-антикодоновое взаимодействие оказывают влияние как модификации самого кодона, так и природа близлежащих оснований, в особенности расположенных с 3′-конца антикодона. Неоднозначное спаривание при взаимодействии кодона и антикодона позволяет митохондриям позвоночных использовать всего 22 тРНК для того, чтобы узнавать все кодоны (вместо обычного для клеток минимума в 31 тРНК); такой эффект достигается в результате изменений митохондриального генетического кода.

Каждая аминокислота распознается своей собственной аминоацил-тРНК-синтетазой, которая одновременно распознает все тРНК, соответствующие данной аминокислоте. Амноацил-тРНК-синтетазы обладают корректирующей активностью, позволяющей тщательно «анализировать» каждую синтезированную аминоацил-тРНК и гидролизовать ошибочные.

Структура аминоацил-тРНК-синтетаз высоковариабельна, но их делят на два класса на основе структуры каталитического домена. Синтетазы связываются с одной из сторон молекулы тРНК, образуя контакты с концами акцепторного стебля и антикодоновой петли; два класса аминоацил-тРНК-синтетаз связываются с молекулой тРНК с противоположных сторон. Относительный вклад акцепторного стебля и антикодонового участка в специфичность взаимодействия между тРНК и ферментом различается для разных тРНК.

Мутации позволяют тРНК прочитывать «не свои» кодоны; большинство таких мутаций возникает в антикодоне. В результате изменения последовательности антикодона тРНК приобретает способность супрессировать некоторые мутации. тРНК, способная узнавать стоп-кодон, является нонсенс-супрессором; тРНК, заменяющая одну аминокислоту на другую, называется миссенс-супрессором. Супрессоры кодонов UAG и UGA более эффективны, чем супрессоры кодона UAA, что объясняется высокой частотой использования UAA в качестве естественного терминирующего кодона. Эффективность всех супрессоров, тем не менее, зависит от контекста каждого индивидуального кодона-мишени.

Сдвиг рамки считывания на один кодон вперед может быть вызван аберрантной тРНК, способной прочитывать «кодоны», состоящие из четырех оснований. Кроме того, сдвиг рамки считывания на один кодон вперед или на один кодон назад может быть вызван «скользкими» последовательностями в мРНК, индуцирующими соскальзывание пептидил-тРНК со своего кодона на перекрывающийся кодон. Такой механизм сдвига рамки требует наличия дополнительной последовательности, вызывающей паузу в работе рибосомы. В ряде случаев сдвиг рамки считывания, зависящий от нуклеотидной последовательности мРНК, является необходимым условием естественного процесса экспрессии гена. «Перепрыгивание» рибосомы сразу через несколько кодонов осуществляется следующим образом: рибосома останавливает трансляцию и скользит вдоль мРНК, удерживая пептидил-тРНК в сайте Р. Затем трансляция возобновляется.

Литература

9.1 Введение

Исследование

Nirenberg, M. W. and Leder, P. (1964). The effect of trinucleotides upon the binding of sRNA to ribosomes. *Science* 145, 1399–1407.

Nirenberg, M. W. and Matthaei, H. J. (1961). The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl. Acad. Sci. USA* 47, 1588–1602.

9.3 Кодон-антикодоновое распознавание допускает неоднозначное спаривание

Исследование

Crick, F. H. C. (1966). Codon-anticodon pairing: the wobble hypothesis. *J. Mol. Biol.* 19, 548–555.

9.4 тРНК образуется из длинного предшественника

Обзор

Hopper, A. K. and Phizicky, E. M. (2003). tRNA transfers to the limelight. *Genes Dev.* 17, 162–180.

9.5 тРНК содержит модифицированные основания

Обзор

Hopper, A. K. and Phizicky, E. M. (2003). tRNA transfers to the limelight. *Genes Dev.* 17, 162–180.

9.6 Модифицированные основания влияют на кодон-антикодоновые взаимодействия

Обзор

Bjork, G. R. (1987). Transfer RNA modification. *Annu. Rev. Biochem.* 56, 263–287.

9.7 Изменения в универсальном генетическом коде носят спорадический характер

Обзор

Fox, T. D. (1987). Natural variation in the genetic code. *Annu. Rev. Genet.* 21, 67–91.

Osawa, S. et al. (1992). Recent evidence for evolution of the genetic code. *Microbiol. Rev.* 56, 229–264.

9.8 Некоторые стоп-кодоны могут встраивать в белок нестандартные аминокислоты

Обзор

Bock, A. (1991). Selenoprotein synthesis: an expansion of the genetic code. *Trends Biochem. Sci.* 16, 463–467.

Ibba, M. and Söll, D. (2004). Aminoacyl-tRNAs: setting the limits of the genetic code. *Genes Dev.* 18, 731–738.

Исследование

Fagegaltier, D., Hubert, N., Yamada, K., Mizutani, T., Carbon, P., and Krol, A. (2000). Characterization of mSelB, a novel mammalian elongation factor for selenoprotein translation. *EMBO J.* 19, 4796–4805.

Hao, B., Gong, W., Ferguson, T. K., James, C. M., Krzycki, J. A., and Chan, M. K. (2002). A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science* 296, 1462–1466.

Srinivasan, G., James, C. M., and Krzycki, J. A. (2002). Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science* 296, 1459–1462.

9.9 Синтетазы присоединяют аминокислоты к тРНК

Обзор

Schimmel, P. (1989). Parameters for the molecular recognition of tRNAs. *Biochemistry* 28, 2747–2759.

9.10 Аминоацил-тРНК-синтетазы делят на 2 класса

Обзор

Schimmel, P. (1987). Aminoacyl-tRNA synthetases: general scheme of structure-function relationships on the polypeptides and recognition of tRNAs. *Annu. Rev. Biochem.* 56, 125–158.

Исследование

Rould, M. A. et al. (1989). Structure of *E. coli* glutaminyl-tRNA synthetase complexed with tRNAGln and ATP at 28 Å resolution. *Science* 246, 1135–1142.

Ruff, M. et al. (1991). Class II aminoacyl tRNA synthetases: crystal structure of yeast aspartyl-tRNA synthetase complexes with $tRNA^{Asp}$. *Science* 252, 1682–1689.

9.11 Синтетазы — ферменты с корректирующей активностью

Обзор

Jakubowski, H. and Goldman, E. (1992). Editing of errors in selection of amino acids for protein synthesis. *Microbiol. Rev.* 56, 412–429.

Исследование

Dock-Bregeon, A., Sankaranarayanan, R., Romby, P., Caillet, J., Springer, M., Rees, B., Francklyn, C. S., Ehresmann, C., and Moras, D. (2000). Transfer RNA-mediated editing in threonyl-tRNA synthetase. The class II solution to the double discrimination problem. *Cell* 103, 877–884.

Hopfield, J. J. (1974). Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc. Natl. Acad. Sci. USA* 71, 4135–4139.

Jakubowski, H. (1990). Proofreading in vivo: editing of homocysteine by methionyl-tRNA synthetase in *E. coli. Proc. Natl. Acad. Sci. USA* 87, 4504–4508.

Nomanbhoy, T. K., Hendrickson, T. L., and Schimmel, P. (1999). Transfer RNA-dependent translocation of misactivated amino acids to prevent errors in protein synthesis. *Mol. Cell* 4, 519–528.

Nureki, O. et al. (1998). Enzyme structure with two catalytic sites for double sieve selection of substrate. *Science* 280, 578–581.

Silvian, L. F., Wang, J., and Steitz, T. A. (1999). Insights into editing from an Ile-tRNA synthetase structure with $tRNA^{Ile}$ and mupirocin. *Science* 285, 1074–1077.

9.14 Супрессорные тРНК могут конкурировать с тРНК дикого типа за связывание с кодонами

Обзор

Atkins, J. F. (1991). Towards a genetic dissection of the basis of triplet decoding, and its natural subversion: programmed reading frameshifts and hops. *Annu. Rev. Genet.* 25, 201–228.

Beier, H. and Grimm, M. (2001). Misreading of termination codons in eukaryotes by natural nonsense suppressor tRNAs. *Nucleic Acids Res.* 29, 4767–4782.

Eggertsson, G. and Soll, D. (1988). Transfer RNA-mediated suppression of termination codons in *E. coli. Microbiol. Rev.* 52, 354–374.

Murgola, E. J. (1985). tRNA, suppression, and the code. *Annu. Rev. Genet.* 19, 57–80.

Normanly, J. and Abelson, J. (1989). Transfer RNA identity. *Annu. Rev. Biochem.* 58, 1029–1049.

Исследование

Hirsh, D. (1971). Tryptophan transfer RNA as the UGA suppressor. *J. Mol. Biol.* 58, 439–458.

Weiner, A. M. and Weber, K. (1973). A single UGA codon functions as a natural termination signal in the coliphage q beta coat protein cistron. *J. Mol. Biol.* 80, 837–855.

9.15 Рибосомы влияют на точность трансляции

Обзор

Kurland, C. G. (1992). Translational accuracy and the fitness of bacteria. *Annu. Rev. Genet.* 26, 29–50.

Ogle, J. M., and Ramakrishnan, V. (2005). Structural insights into translational fidelity. *Annu. Rev. Biochem.* 74, 129–177.

Ramakrishnan, V. (2002). Ribosome structure and the mechanism of translation. *Cell* 108, 557–572.

Исследование

Carter, A. P., Clemons, W. M., Brodersen, D. E., Morgan-Warren, R. J., Wimberly, B. T., and Ramakrishnan, V. (2000). Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* 407, 340–348.

Ogle, J. M., Brodersen, D. E., Clemons, W. M., Tarry, M. J., Carter, A. P., and Ramakrishnan, V. (2001). Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science* 292, 897–902.

9.17 Сдвиг рамок считывания происходит на «скользких» последовательностях

Обзор

Farabaugh, P. J. (1995). Programmed translational frameshifting. *Microbiol. Rev.* 60, 103–134.

Farabaugh, P. J. and Bjorkk, G. R. (1999). How translational accuracy influences reading frame maintenance. *EMBO J.* 18, 1427–1434.

Gesteland, R. F. and Atkins, J. F. (1996). Recoding: dynamic reprogramming of translation. *Annu. Rev. Biochem.* 65, 741–768.

Исследования

Jacks, T., Power, M. D., Masiarz, F. R., Luciw, P. A., Barr, P. J., and Varmus, H. E. (1988). Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature* 331, 280–283.

9.18 «Перепрыгивание» через основания осуществляется пермещением рибосомы

Обзор

Herr, A. J., Atkins, J. F., and Gesteland, R. F. (2000). Coupling of open reading frames by translational bypassing. *Annu. Rev. Biochem.* 69, 343–372.

Исследование

Gallant, J. A. and Lindsley, D. (1998). Ribosomes can slide over and beyond "hungry" codons, resuming protein chain elongation many nucleotides downstream. *Proc. Natl. Acad. Sci. USA* 95, 13771–13776.

Huang, W. M., Ao, S. Z., Casjens, S., Orlandi, R., Zeikus, R., Weiss, R., Winge, D., and Fang, M. (1988). A persistent untranslated sequence within bacteriophage T4 DNA topoisomerase gene 60. *Science* 239, 1005–1012.

Локализация белка

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

10.1 Введение

10.2 Транспорт белков через мембрану требует участия специального белкового аппарата

- Транспорт белков через мембрану происходит при участии специализированных белковых структур, входящих в состав мембраны.
- Синтезируемый белок напрямую взаимодействует с белок-транспортирующим комплексом на мембране ЭПР, митохондрий или хлоропластов, но требует участия дополнительных белков-переносчиков при транспорте в пероксисомы.
- Для транспорта белков в ядро требуется участие гораздо более сложного белкового аппарата.

10.3 Транслокация может быть как пост-, так и котрансляционной

- Белки, импортируемые в цитоплазматические органеллы, синтезируются в цитозоле на свободных рибосомах.
- Белки, импортируемые в систему ЭПР–комплекс Гольджи, синтезируются на рибосомах, ассоциированных с мембраной ЭПР.
- Белки связываются с мембраной ЭПР посредством специальных аминокислотных последовательностей, называемых сигнальными последовательностями.
- Сигнальными чаще всего являются первые несколько аминокислот белка (лидерный участок) на N-конце молекулы белка.
- N-концевые сигнальные последовательности, как правило, отщепляются во время прохождения белка через мембрану.

10.4 Шапероны обеспечивают формирование правильной пространственной структуры белка

- Белки, способные самопроизвольно принимать нужную конформацию, называются самособирающимися белками.
- Иногда белки могут принимать несколько альтернативных конформаций.
- Шапероны направляют укладку полипептидной цепи белка по единственно правильному пути, исключая возможные альтернативные варианты.
- Шапероны предотвращают формирование неправильных структур, взаимодействуя с еще не свернувшимся белком и направляя его сворачивание.

10.5 Шапероны необходимы и для вновь синтезированных, и для денатурированных белков

- Шапероны действуют на синтезируемые, транспортируемые через мембрану и на денатурированные белки.
- Hsp (белок теплового шока) 70 и несколько связанных с ним белков формируют основной класс шаперонов, взаимодействующих с большим числом белков-мишеней.
- Шаперонины I и II групп представляют собой большие олигомерные белковые комплексы, взаимодействующие с белками-мишенями путем их заключения в своей внутренней полости.
- Hsp90 — специализированный шаперон, действующий на белки сигнальных каскадов.

10.6 Белки семейства Hsp70 встречаются во всех компартментах клетки

- Белки — члены семейства Hsp70 — обнаружены в цитоплазме, ЭПР, митохондриях и хлоропластах.
- Hsp70 — шаперон, действующий на белок-мишень в комплексе с DnaJ и GrpE.

10.7 Сигнальные последовательности инициируют транслокацию

- Белки попадают в систему ЭПР только котрансляционно.
- За ассоциацию с мембраной ЭПР ответственны сигнальные последовательности синтезируемых белков.

10.8 Сигнальные последовательности взаимодействуют с сигнал-узнающими частицами (SRP)

- Сигнальные последовательности связываются с SRP (сигнал-узнающими частицами).
- Связывание SRP с сигнальной последовательностью белка вызывает остановку белкового синтеза.
- Синтез белка возобновляется, когда SRP связывается с SRP-рецептором на мембране ЭПР.
- Расположенная на внутренней поверхности мембраны сигнальная пептидаза отщепляет сигнальные последовательности от транспортируемого белка.

10.9 SRP взаимодействует со своим рецептором

- SRP представляет собой комплекс из 7S РНК и шести белков.

- Бактериальным эквивалентом SRP является комплекс из 4,5S РНК и двух белков.
- SRP-рецептор представляет собой димер.
- После взаимодействия с рецептором SRP диссоциирует от него в результате гидролиза ГТФ.

10.10 Транслокон формирует пору

- Тримерный комплекс Sec61 формирует канал для транспорта белков через мембрану.
- Транспортируемый через мембрану белок поступает в транслокон непосредственно с рибосомы (без контакта с цитоплазмой).

10.11 Для транслокации необходимы дополнительные включения в транслокон, а иногда и храповик в ЭПР

- Рибосома, SRP и SRP-рецептор создают условия для попадания насцентного белка в транслокон.
- Белки, попадающие в транслокон посттранляционно, требуют наличия дополнительных компонентов в цитоплазме и белка BiP в ЭПР.
- BiP — это храповик, предотвращающий движение белка в обратную сторону.

10.12 Для деградации белки возвращаются в цитозоль путем обратной транслокации

- Транслокон Sec61 может использоваться для обратного переноса белков из ЭПР в цитоплазму.

10.13 Белки заякорены в мембране при помощи гидрофобных регионов

- У мембранных белков I группы N-конец молекулы расположен на внутренней стороне мембраны, а белки II группы имеют противоположную ориентацию.
- Некоторые белки имеют несколько мембрансвязанных доменов.

10.14 Якорные последовательности определяют ориентацию белка в мембране

- Якорная последовательность останавливает движение белка через транслокон. Как правило, она расположена на C-конце молекулы, приводя к образованию мембранного белка I группы с N-концом, прошедшим сквозь мембрану и экспонированным на внутренней ее стороне.
- Комбинированная сигнально-якорная последовательность служит одновременно и для введения белка сквозь мембранный канал, и для заякоривания в мембране вошедшего участка молекулы. Поскольку заякоривание идет с внутренней стороны мембраны, мембранный белок приобретает ориентацию белка II группы с N-концом, экспонированным в цитоплазму.

10.15 Как белки встраиваются в мембраны?

- Переход трасмембранных доменов из транслокона в липидный бислой запускается в результате взаимодействия трансмембранного региона с транслоконом.

10.16 Посттрансляционная интеграция белков в мембрану зависит от лидерных последовательностей

- N-концевая лидерная последовательность несет информацию, необходимую для ассоциации белка с мембранами митохондрий или хлоропластов.

10.17 Серия лидерных участков определяет детальное местоположение белка внутри органеллы

- N-терминальная часть лидерной последовательности определяет попадание белка в матрикс митохондрий или внутреннее пространство хлоропластов.
- Последующие участки лидерного пептида определяют дальнейшее встраивание белка в мембрану органеллы либо его попадание в межмембранное пространство.
- Лидерные последовательности отщепляются от белка одна за другой.

10.18 Внешняя и внутренняя мембраны митохондрий содержат разные транслоконы

- Транспорт белков через внешнюю и внутреннюю мембраны осуществляется при участии разных рецепторных комплексов.
- Расположенный на внешней мембране комплекс TOM представляет собой крупную белковую структуру, направляющую транспортируемые белки в канал Tom40 при помощи одного из двух субкомплексов.
- В зависимости от того, куда должен попасть транспортируемый белок (на внутреннюю мембрану или в матрикс), на внутренней мембране митохондрий для транспорта используются разные TIM-комплексы.
- Белки переходят от комплекса TOM непосредственно к комплексу TIM.

10.19 Пероксисомы обладают собственной системой транслокации белков

- Белки импортируются в пероксисомы в виде молекул с уже правильной пространственной структурой.
- Они имеют либо последовательность PTS1 на C-конце молекулы либо последовательность PTS2 на N-конце.
- Рецептор Pex5p связывает последовательность PTS1, рецептор Pex7p связывает последовательность PTS2.
- Рецепторы представляют собой цитоплазматические белки, перемещающиеся в пероксисому в комплексе с транспортируемым белком и затем возвращающиеся в обратно в цитоплазму.

10.20 Бактерии используют как котрансляционную, так и посттрансляционную системы транспорта белков

- Для транспорта бактериальных белков на мембрану или сквозь нее используются оба механизма транслокации: котрансляционный и посттрансляционный.

10.21 Sec-система транспортирует белки на внутреннюю мембрану бактериальной оболочки и сквозь нее

- Бактериальный транслокон SecYEG соответствует транслокону Sec61 у эукариот.
- В процесс попадания транспортируемого белка в транслокон вовлечены различные шапероны.

10.22 Sec-независимая система транспорта белков *E. coli*

- Системы для транслокации белков через клеточную мембрану *E. coli* сходны с системами, характерными для органелл.
- Одна из систем позволяет определенным белкам встраиваться в мембрану без участия аппарата транслокации.
- YidC гомологичен митохондриальной системе переноса белков во внутреннюю мембрану.
- Tat-система транспортирует белки с двойным аргининовым мотивом в периплазматическое пространство.

10.23 Резюме

10.1 Введение

Синтез белка имеет два типа локализации в клетке:

- Подавляющее большинство белков синтезируется цитоплазматическими рибосомами.
- Небольшая часть белков синтезируется рибосомами внутри органелл (митохондрий или хлоропластов).

В зависимости от локализации, синтезирующие в цитоплазме рибосомы делят на свободные и мембрансвязанные. На РИС. 10.1 изображен внутриклеточный транспорт белков с указанием возможных конечных пунктов назначения:

- Цитоплазматические (или «растворенные») белки не связаны с какими-либо органеллами. Они синтезируются в цитоплазме и остаются там же, функционируя как индивидуальные каталитические центры и действуя на присутствующие в цитоплазме метаболиты.
- Макромолекулярные структуры могут располагаться в определенных участках цитоплазмы; например, местоположение центриолей ассоциировано с регионами, где впоследствии возникают полюса веретена деления.
- Ядерные белки транспортируются от мест своего синтеза в цитоплазме через ядерную оболочку в ядро.
- Большинство белков цитоплазматических органелл синтезируются в цитоплазме и специфично транспортируются к мембране (и сквозь нее) митохондрий, пероксисом или (в растительных клетках) хлоропластов. (Белки, синтезируемые внутри органеллы, остаются внутри нее.)

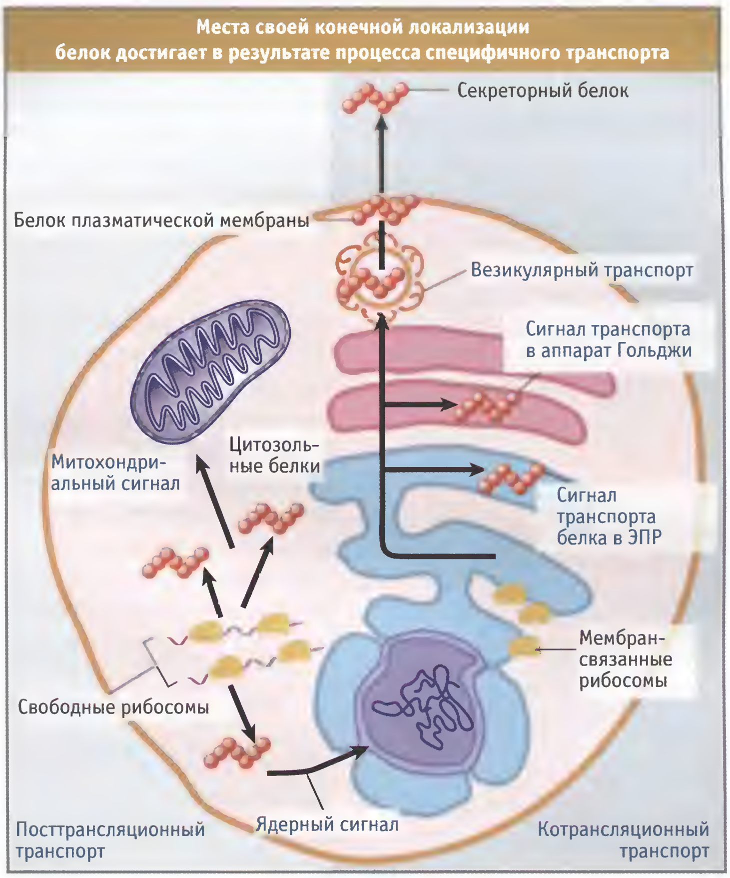


РИС. 10.1 Белки, транспорт которых осуществляется посттрансляционно, после синтеза на свободных рибосомах высвобождаются в цитозоль. Некоторые из этих белков обладают сигналами, направляющими их в органеллы (ядро, митохондрии). Белки, транспорт которых осуществляется котрансляционно, ассоциируют с мембраной ЭПР во время синтеза. Рибосомы, синтезирующие такие белки, называют мембрансвязанными. Белки поступают в эндоплазматический ретикулум, затем через аппарат Гольджи направляются к плазматической мембране и транспортируются сквозь нее (в том случае, если не обладают сигналом, определяющим остановку движения белка на каком-либо из перечисленных шагов). Конечной точкой транспорта белков могут быть и другие органеллы, такие как эндосомы и лизосомы

- В цитоплазме расположен ряд мембранных компартментов, таких как эндоплазматический ретикулум (ЭПР), аппарат Гольджи, эндосомы и лизосомы. Иногда их называют ретикулоэндотелиальной системой. Белки, обнаруживаемые внутри этих компартментов, сначала проходят сквозь мембрану ЭПР, после чего направляются к конечному месту своей локализации с помощью транспортной системы аппарата Гольджи.
- Секретируемые белки транспортируются к плазматической мембране и выбрасываются во внеклеточное пространство. Начальные этапы их синтеза не отличаются от синтеза белков, ассоциированных с ретикулоэндотелиальной системой, однако вместо того, чтобы задержаться в каком-либо клеточном компартменте, они проходят весь транспортный путь насквозь.

10.2 Транспорт белков через мембрану требует участия специального белкового аппарата

Ключевые понятия:

- Транспорт белков через мембрану происходит при участии специализированных белковых структур, входящих в состав мембраны.
- Синтезируемый белок напрямую взаимодействует с белок-транспортирующим комплексом на мембране ЭПР, митохондрий или хлоропластов, но требует участия дополнительных белков-переносчиков при транспорте в пероксисомы.
- Для транспорта белков в ядро требуется участие гораздо более сложного белкового аппарата.

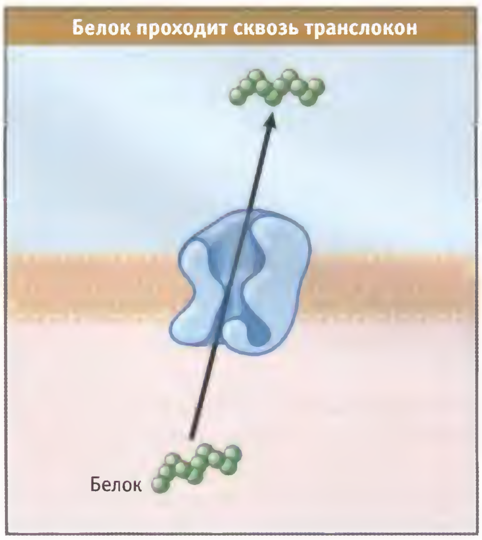


РИС. 10.2 Белки транспортируются внутрь ЭПР или в митохондрии путем связывания с транслоконом, осуществляющим их перенос через мембрану

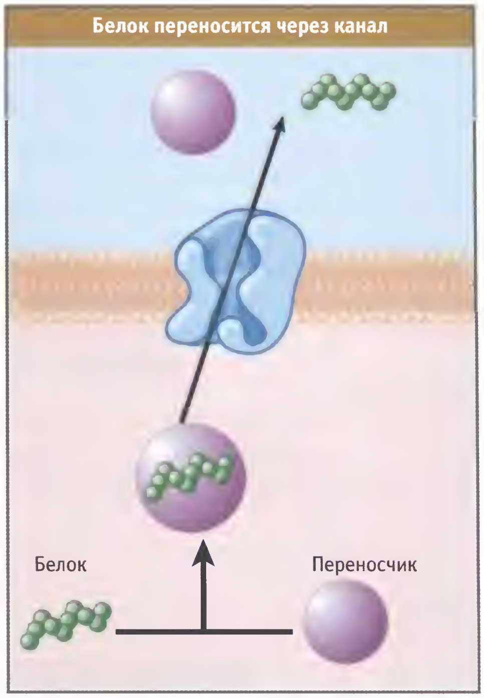


РИС. 10.3 Белки транспортируются в пероксисомы с помощью белка-переносчика, который связывает их в цитозоле, переносит через мембранный канал и высвобождает на другой стороне мембраны

Процесс интеграции белков в мембрану или их переноса через мембрану называют **транслокацией.** Вопрос выбора между встраиванием белка в мембрану и транспортом сквозь нее встает каждый раз при прохождении белка через мембрану. Белок, как правило, обладает гидрофильной поверхностью, в то время как мембрана является гидрофобной. Подобно маслу и воде, гидрофильные белки и мембраны не смешиваются. Решение проблемы заключается в создании специальной структуры в мембране, сквозь которую мог бы пройти гидрофильный белок. Известно три различных типа таких структур.

Эндоплазматический ретикулум, митохондрии и хлоропласты содержат в составе своих мембран структуры, позволяющие белкам проходить сквозь мембрану (не контактируя при этом с гидрофобными липидами). На РИС. 10.2 изображено, как транспортируемый белок связывается с такой структурой, переносится ею сквозь мембрану и высвобождается на другой стороне.

Пероксисомы тоже имеют аналогичные структуры в своих мембранах, однако транспортируемые белки не связываются непосредственно с ними. На РИС. 10.3 показано, как транспортируемый белок связывается с белком-переносчиком в цитоплазме, переносится с его помощью внутрь пероксисомы через мембранный канал, после чего диссоциирует от преносчика.

Для транспорта белков в ядро требуется участие гораздо более сложного механизма, ключевую роль в ко-

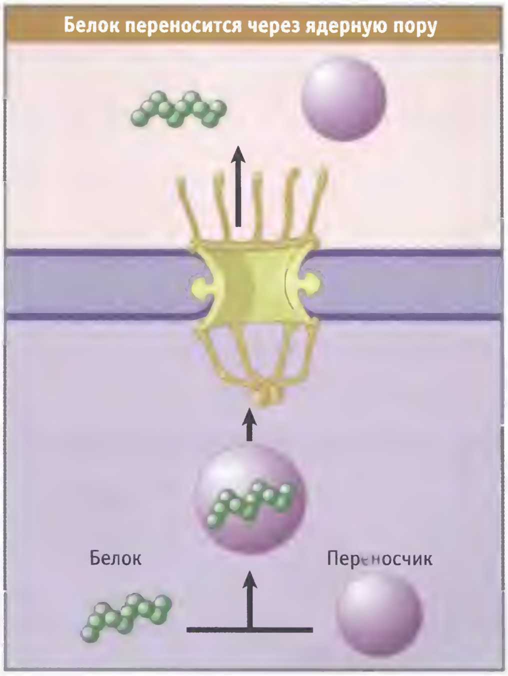


РИС. 10.4 Белки проникают в ядро, проходя через очень большие ядерные поры. Транспортный аппарат существует отдельно от самой поры и представлен белками-переносчиками, доставляющими ядерные белки внутрь ядра

тором играет ядерная пора. Как показано на РИС. 10.4, ядерная пора создает условия, позволяющие субстрату проникнуть в ядро (или выйти из него), однако сама по себе она не связывает белок и не переносит его. Для этого существуют белки-переносчики, связывающие субстрат и транспортирующие его сквозь ядерную пору на другую сторону ядерной оболочки.

10.3 Транслокация может быть как пост-, так и котрансляционной

Ключевые понятия:

- Белки, импортируемые в цитоплазматические органеллы, синтезируются в цитозоле на свободных рибосомах.
- Белки, импортируемые в систему ЭПР–комплекс Гольджи, синтезируются на рибосомах, ассоциированных с мембраной ЭПР.
- Белки связываются с мембраной ЭПР посредством специальных аминокислотных последовательностей, называемых сигнальными последовательностями.
- Сигнальными чаще всего являются первые несколько аминокислот белка (лидерный участок) на N-конце молекулы белка.
- N-концевые сигнальные последовательности, как правило, отщепляются во время прохождения белка через мембрану.

Существует два пути, с помощью которых белки устанавливают первичный контакт с мембраной:

- Белок может связываться с аппаратом транслокации во время своего синтеза на рибосоме. Такой тип переноса белков называют **котрансляционной транслокацией.**
- Другой путь заключается в синтезе полного белка на рибосоме, после чего синтезированный белок диффундирует к мембране и связывается с аппаратом транслокации. Такой тип переноса белков называют **посттрансляционной транслокацией.**

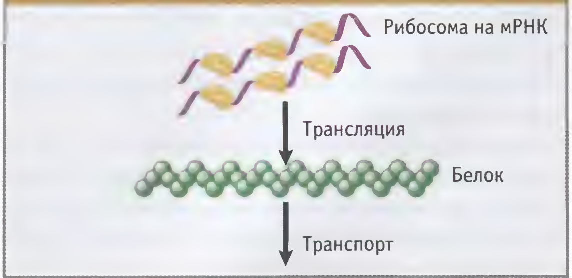
Следовательно, местоположение работающей рибосомы будет зависеть от того, ассоциирует белок с мембраной котрансляционно (в ходе синтеза) или посттрансляционно:

- Котрансляционная транслокация присуща белкам, транспортируемым в эндоплазматический ретикулум (ЭПР). В процессе синтеза такого белка рибосома находится на поверхности ЭПР. Поскольку рибосомы на время котрансляционной транслокации ассоциированы с мембраной ЭПР, они обнаруживаются в мембранной фракции клеток (иногда такие рибосомы называют мембрансвязанными).
- Другие рибосомы находятся в цитоплазме. Поскольку они не ассоциированы ни с какими органеллами и фракционируются отдельно от клеточных мембран, иногда их называют «свободными рибосомами». Свободные рибосомы синтезируют все белки клетки, кроме котрансляционно транслоцируемых в ЭПР белков. После завершения синтеза на свободных рибосомах белки высвобождаются в цитоплазму. Часть этих белков так и остается в цитоплазме в свободной форме; другие ассоциируют с макромолекулярными цитоплазматическими структурами (такими, как микрофиламенты, микротрубочки, центриоли, и т. д.), транспортируются в ядро или попадают в органеллы в результате посттрансляционной транслокации.

Для того чтобы связаться с мембраной (или любой другой структурой), белок должен нести соответствующий сигнал. Как правило, для этого используется определенная аминокислотная последовательность, узнаваемая системой транслокации (или же встраивание белка в специальную макромолекулярную структуру).

На РИС. 10.5 перечислены некоторые сигналы, используемые белками, синтезируемыми на свободных рибосомах. Транспорт белка в ядро опосредован набором разнообразных, сравнительно коротких, последовательностей, входящих в состав самого белка. Эти «сигналы ядерной локализации» обеспечивают прохождение белка через ядерные поры. Сигналом, определяющим транспорт белка в пероксисому, является очень короткая С-концевая последовательность. Митохондриальные белки и белки хлоропластов синтезируются свободными рибосомами. Уже после высвобождения в цитоплазму они ассоциируют с мембранами органелл посредством N-концевых последовательностей (длиной около 25 аминокислот), узнаваемых белками оболочки органеллы.

Место будущей локализации белка определяется короткими сигнальными последовательностями



Органелла	Место расположения сигнала	Природа сигнала	Длина сигнальной последовательности
Митохондрия	N-концевой	Амфипатическая спираль	12–30
Хлоропласт	N-концевой	Заряд	>25
Ядро	Внутренний	Одинарный или двойной	4–9
Пероксисома	С-концевой	Короткий пептид	3–4

РИС. 10.5 Белки, синтезируемые свободными рибосомами в цитозоле, направляются к своему конечному месту назначения при участии коротких сигнальных последовательностей

Белки могут проникнуть в ЭПР только во время синтеза



РИС. 10.6 Белки проникают в систему ЭПР–комплекс Гольджи, только будучи ассоциироваными с мембраной эндоплазматического ретикулума во время синтеза

Белки ретикулоэндотелиальной системы попадают в ЭПР в процессе синтеза. Принципы котрансляционной транслокации изображены на РИС. 10.6. Синтезируемый белок сам отвечает за узнавание транслокационного аппарата, что является важной характеристикой этой системы. Механизм такого транспорта основан на том, что сигнал котрансляционной транслокации является частью белка, синтезируемой в первую очередь. Действительно, сигнальные последовательности котрансляционной транслокации расположены обычно на N-конце молекулы белка.

В белках, использующих N-концевые сигналы для котрансляционного транспорта в ЭПР и посттрансляционного транспорта в митохондрии и хлоропласты, обнаружена одна общая черта. N-концевая или **лидерная** последовательность не обнаруживается в составе молекулы зрелого белка. Белок, несущий такую последовательность, называется **пре-белком** и представляет собой короткоживущий предшественник зрелого белка. Лидерная последовательность отщепляется от белка во время процесса транслокации через мембрану ЭПР.

10.4 Шапероны обеспечивают формирование правильной пространственной структуры белка

Ключевые понятия:

- Белки, способные самопроизвольно принимать нужную конформацию, называются самособирающимися белками.
- Иногда белки могут принимать несколько альтернативных конформаций.
- Шапероны направляют укладку полипептидной цепи белка по единственно правильному пути, исключая возможные альтернативные варианты.
- Шапероны предотвращают формирование неправильных структур, взаимодействуя с еще не свернувшимся белком и направляя его сворачивание.

Некоторые белки способны самопроизвольно принимать конформацию, свойственную их зрелому состоянию. Проверить эту способность можно, подвергнув белок денатурации и проследив, сможет ли он заново ренатурировать в свою активную форму. Такое свойство белков называют **самосборкой**. Самособирающийся белок способен сворачиваться в активную форму из неактивных конформаций (включая неверные конформации, реализуемые при его синтезе на рибосоме). Эта способность определяется внутренними взаимодействиями в молекуле белка, направлеными на формирование активной конформации. Классическим примером самособирающихся белков являются рибонуклеазы. В 1970 году было показано, что после денатурации этот фермент может восстанавливать свою правильную конформацию *in vitro*. Позднее процесс сворачивания был детально описан для нескольких небольших белков.

Однако некоторые белки не способны к самостоятельному формированию корректной пространственной структуры, и белок принимает «неверную» стабильную конформацию, определяемую альтернативными внутримолекулярными взаимодействиями. Белки такого типа не обладают способностью к самосборке. Формирование активной конформации в этом случае требует участия **шаперонов.**

Фолдинг (сворачивание) белка обычно происходит посредством взаимодействия между несколькими контактными поверхностями. Как правило, это поверхности, несущие выступающие наружу гидрофобные радикалы, формирующие гидрофобное ядро белка. В отсутствие шаперонов гидрофобность этих поверхностей становится причиной неправильных взаимодействий и, как следствие, неправильной укладки полипептидной цепи.

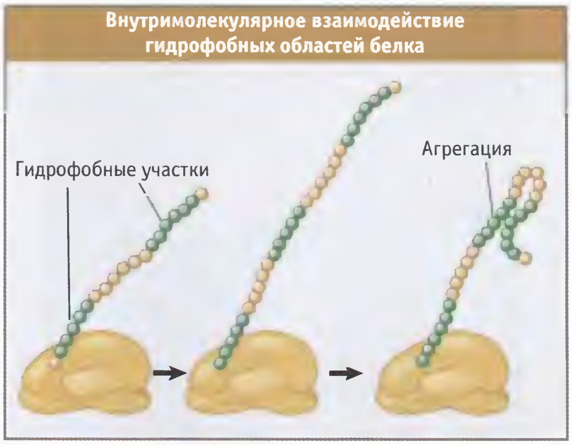


РИС. 10.7 Гидрофобные регионы одной полипептидной цепи синтезирующегося или денатурированного белка взаимодействуют друг с другом

На РИС. 10.7 изображено, что может произойти в данном случае. По мере синтеза белка на рибосоме, гидрофобные участки, покидая зону влияния рибосомы, в водной среде стремятся агрегировать с другими гидрофобными участками. Поскольку такие взаимодействия возникают случайно, они могут и не соответствовать активной конформации белка.

Шапероны — белки, определяющие правильную конформацию белков-мишеней, за счет влияния на процесс их сворачивания. Эффект от действия шаперонов достигается благодаря связыванию ими экспонируемых во время синтеза белка-мишени контактных поверхностей, тем самым предотвращая взаимодействие этих поверхностей с другими регионами белка и формирование неправильной конформации. Функция шаперонов заключается скорее в предотвращении формирования неправильных структур, чем в создании правильной. На РИС. 10.8 показан пример того, как шаперон связывает гидрофобный участок полипептидной цепи, обеспечивая взаимодействие, которое было бы невозможно в его отсутствие (сравните с рис. 10.7).

Неправильная пространственная структура может формироваться в результате ошибочного фолдинга одного белка либо в результате взаимодействия белков друг с другом. Концентрация белков в цитоплазме довольно высока, что может значительно увеличивать вероятность многих взаимодействий в такой «макромолекулярной каше» в сравнении с ситуацией *in vitro*. Скопление белков способствует их агрегации, однако шапероны помогают предотвратить нежелательные эффекты. Таким образом, роль шаперонов состоит в защите белков от агрегации и в обеспечении концентрационно-независимого фолдинга белков в цитоплазме.

Точно неизвестно, какая часть белков обладает способностью к самосборке, а какая требует участия шаперонов. (Заметим, что способность белка к самосборке *in vitro* не означает, что он обладает этой способностью и *in vivo* вследствие огромной разницы в условиях протекания реакции. Возможно, для многих белков, самособирающихся в экспериментальных условиях, сворачивание в клетке идет при участии шаперонов. Однако на основании данных *in vitro* можно разделить белки на принципиально способные к самосборке, и белки, требующие обязательного участия шаперонов для формирования правильной пространственной структуры.)

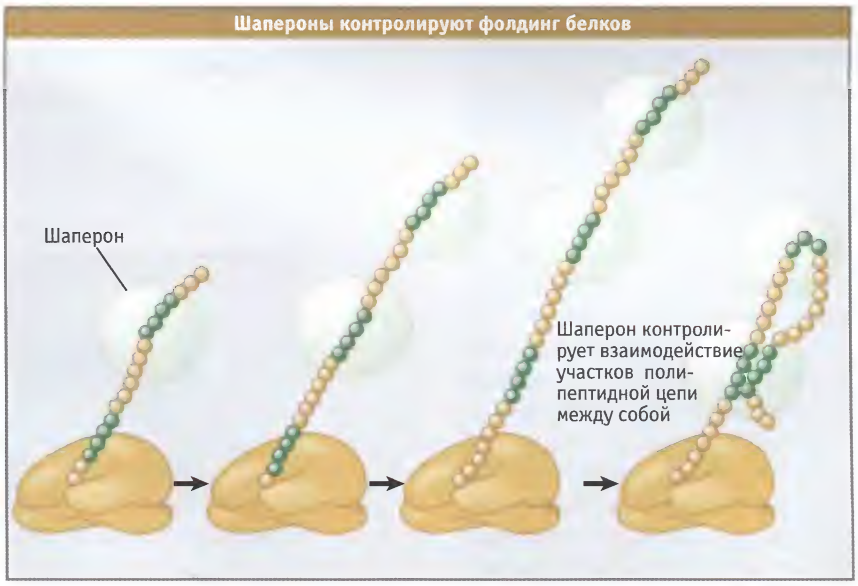


РИС. 10.8 Участки полипептидной цепи, взаимодействующие между собой во время синтеза, связываются шаперонами. Тем самым предотвращается их случайная агрегация и исключается возможность образования ошибочной конформации

10.5 Шапероны необходимы и для вновь синтезированных, и для денатурированных белков

Ключевые понятия:

- Шапероны действуют на синтезируемые, транспортируемые через мембрану, и на денатурированные белки.
- Hsp (белок теплового шока) 70 и несколько связанных с ним белков формируют основной класс шаперонов, взаимодействующих с большим числом белков-мишеней.
- Шаперонины I и II групп представляют собой большие олигомерные белковые комплексы, взаимодействующие с белками-мишенями путем их заключения в своей внутренней полости.
- Hsp90 — специализированный шаперон, действующий на белки сигнальных каскадов.

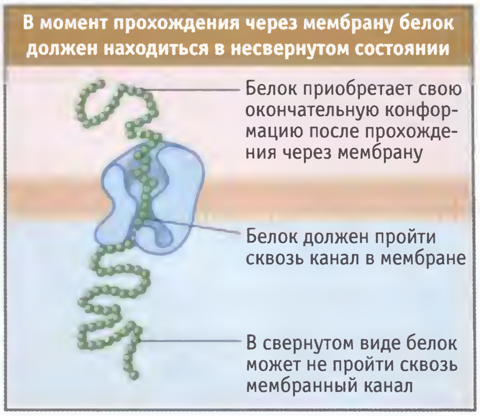


РИС. 10.9 Форма белка при прохождении его через мембрану определяется узким транспортным каналом

Способность шаперонов узнавать неверную конформацию белка позволяет им выполнять две взаимосвязанные функции, касающиеся структуры белка:

- Выходящий по мере синтеза (на рибосоме) в цитоплазму белок представляет собой несвернутую форму. Укладка полипептида идет по мере того, как синтезированные участки последовательности выходят из рибосомы и начинают взаимодействовать друг с другом. Шапероны влияют на это взаимодействие, контролируя доступность контактных поверхностей. Этот процесс называют первичным фолдингом белка.
- Если белок денатурирует, гидрофобные регионы снова становятся открытыми для взаимодействия (в любой комбинации). Эти взаимодействия напоминают процессы, происходящие при случайном (ошибочном) фолдинге белка во время синтеза. Такие «испортившиеся» белки могут распознаваться шаперонами как обладающие неправильной пространственной структурой и либо подвергаются ренатурации, либо устраняются путем деградации.

Кроме того, шапероны могут быть задействованы в формировании олигомерных структур и участвовать в транспорте белков через мембраны. Важным моментом в обеспечении трансмембранного переноса белков является контроль (через задержку) их фолдинга. Как показано на **РИС. 10.9**, геометрия трансмембранного переноса требует поддержания белка в несвернутом состоянии до тех пор, пока он не пройдет сквозь мембрану: свернутый белок может быть попросту слишком большим для того, чтобы пройти сквозь мембранный канал. Шапероны способны предотвращать приобретение белком конформации, мешающей его транспорту сквозь мембрану; поэтому их роль сводится, главным образом, к поддержанию белка в несвернутом пластичном состоянии. После того как белок прошел сквозь мембрану, для приобретения конформации, соответствующей его зрелой форме, может потребоваться участие других шаперонов (точно так же, как цитозольные белки требуют участия шаперонов во время синтеза на рибосоме). Состояние белка в момент прохождения сквозь мембрану, видимо, напоминает его состояние в момент, когда он покидает рибосому во время синтеза — молекула и в том, и в другом случае вытягивается, приобретая более или менее линейную форму.

На данный момент две основные группы шаперонов хорошо изучены. Они влияют на фолдинг белков посредством механизмов двух разных типов:

- На **РИС. 10.10** показано, как *система Hsp70 (от heat shock protein — белок теплового шока), состоящая из нескольких отдельных белков, связывается с субстратом, сворачивание которого контролирует*. Hsp70 узнает только что синтезированные белки или белки, прошедшие через мембрану (а также белки, денатурированные во время стресса). В основном эта система контролирует взаимодействия между экспонированными наружу реакционноспособными регионами белка, заставляя белок принимать правильную конформацию *in situ*. Система в целом включает компоненты — Hsp70, Hsp40 и GrpE. Общее название системы (Hsp70) отражает первичную идентификацию Hsp70 как белка, синтез которого индуцируется тепловым шоком. Белки Hsp70 и Hsp40 связываются с белком-субстратом обособленно. С помощью энергии гидролиза АТФ они изме-

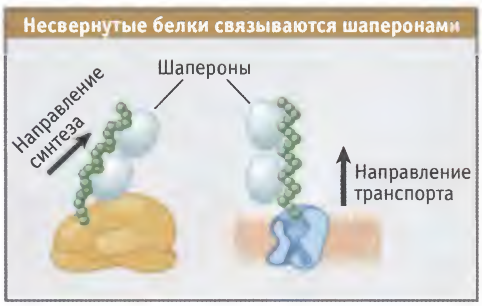


РИС. 10.10 Только что синтезированные рибосомой белки или белки, прошедшие сквозь мембрану, находятся в несвернутом состоянии. Связываясь с такими белками, шапероны не дают им свернуться в неправильную структуру



РИС. 10.11 Шаперонины формируют большой олигомерный комплекс, внутри которого располагаются несвернутые белки

няют структуру субстратного белка. Hsp70 и Hsp40 работают совместно с фактором обмена, восстанавливающим АТФ из АДФ.

- На РИС. 10.11 изображена *система шаперонинов, представленная большим олигомерным комплексом (изображенным в виде цилиндра).* Этот комплекс формирует структуру, внутрь которой помещаются несвернутые белки, и уже внутренняя среда комплекса контролирует сворачивание белка. Различают два типа шаперониновых систем: GroEL/GroES (обнаружена во всех классах живых организмов) и TRiC (присутствует в цитоплазме эукариотических клеток).

Различают два основных типа систем шаперонов	
Система	Структура/функции
Индивидуальные шапероны	
Система Hsp70	
Hsp70 (DnaK)	АТФаза
Hsp40 (DnaJ)	Активатор АТФазы
GrpE (GrpE)	Фактор нуклеотидного обмена
Hsp90	Действует на белки, вовлеченные в сигнальные каскады
Олигомерные структуры (шаперонины)	
Группа I	
Hsp60 (GroEL)	Образует два гептамерных кольца
Hsp10 (GroES)	Формирует «крышку»
Группа II	
TRiC	Образует два октамерных кольца

РИС. 10.12 Шапероны эукариот имеют свои аналоги у бактерий (названия приведены в круглых скобках)

Компоненты всех этих систем перечислены на РИС. 10.12. Система Hsp70 и две шаперониновые системы действуют на множество различных белков. Еще одна система — белок Hsp90 — функционирует совместно с Hsp70, однако ее действие направлено на белки, вовлеченные в передачу внутриклеточных сигналов (на рецепторы стероидных гормонов и сигнальные киназы). Основной функцией системы Hsp90 является поддержание белков-мишеней в подходящей конформации, в то время как они стабилизируются, взаимодействуя с другими компонентами сигнального каскада. Причина, по которой многие из шаперонов носят название Hsp (heat shock protein — белок теплового шока), заключается в том, что повышение температуры вызывает продукцию этих белков, поскольку основная их функция состоит в минимизации разрушительного эффекта тепловой денатурации белков. Многие из шаперонов впервые были открыты (и названы) именно как компоненты ответа на тепловой шок.

10.6 Белки семейства Hsp70 встречаются во всех компартментах клетки

Ключевые понятия:

- Белки — члены семейства Hsp70 — обнаружены в цитоплазме, ЭПР, митохондриях и хлоропластах.
- Hsp70 — шаперон, действующий на белок-мишень в комплексе с DnaJ и GrpE.

Семейство белков Hsp70 обнаружено в бактериях, цитоплазме эукариотических клеток, в ЭПР, хлоропластах и митохондриях. Типичный Hsp70 имеет два домена: N-концевой домен, обладающий АТФ-азной активностью, и С-концевой домен, необходимый для взаимодействия с субстратным полипептидом. В комплексе с АТФ Hsp70 быстро связывает и высвобождает субстрат; если же Hsp70 связан с АДФ, реакция протекает медленно. Переход между АТФ/АДФ состояниями регулируется двумя другими белками: Hsp40 (DnaJ) и GrpE.

Как показано на РИС. 10.13, Hsp40 (DnaJ) первым связывается с синтезирующимся белком (в момент выхода его N-конца на пределы рибосомы). Hsp40 содержит регион, названный J-доменом (отсюда название DnaJ), взаимодействующий с Hsp70. Hsp70 (DnaK) связывается одновременно и с Hsp40, и с несвернутым полипептидом. В результате с субстратным белком оказываются связаны два взаимодействующих между собой шаперона. J домен определяет специфичность парного взаимодействия и отвечает за то, чтобы определенный Hsp40 выбрал себе подходящего партнера среди семейства белков Hsp70.

Взаимодействие Hsp70 (DnaK) с Hsp40 (DnaJ) стимулируется АТФ-азной активностью Hsp70. АДФ-связанная форма такого комплекса остается ассоциированной с субстратным белком до тех пор, пока GrpE не вытеснит

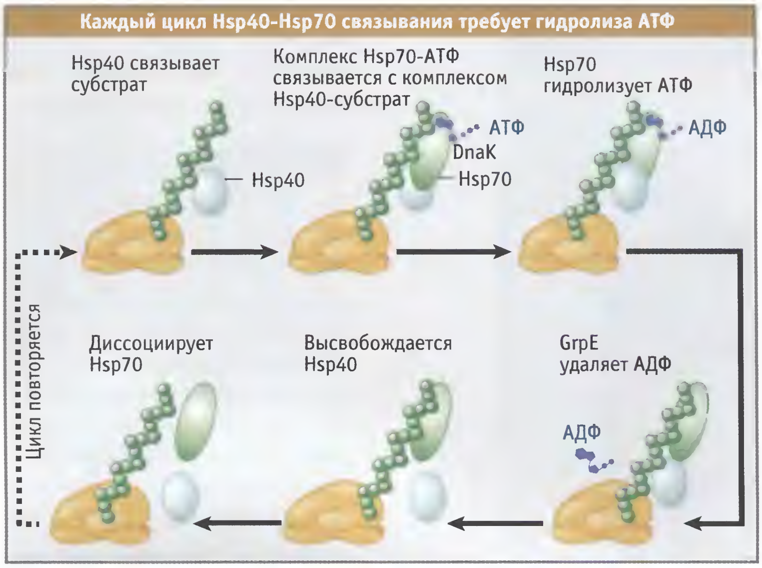


РИС. 10.13 Hsp40 связывает субстратный белок, а затем Hsp70. Гидролиз АТФ запускает конформационные изменения. GrpE вытесняет АДФ, что приводит к высвобождению шаперонов. Фолдинг белка осуществляется посредством множества циклов ассоциации и диссоциации шаперонов

АДФ из комплекса. Это приводит к отсоединению Hsp40, за которым следует и диссоциация Hsp70. Свободный Hsp70 связывает новую молекулу АТФ и цикл повторяется. GrpE (или его аналоги) обнаружены только у бактерий, в митохондриях и в хлоропластах; во всех прочих случаях реакция диссоциации Hsp70 связана с гидролизом АТФ более сложным путем.

Сворачивание белка осуществляется посредством многократного повторения циклов ассоциации и диссоциации шаперонов. По мере того как длина полипептидной цепи увеличивается, Hsp70 (DnaK) диссоциирует с одного сайта связывания и реассоциирует с другим, таким образом позволяя участкам субстратного белка правильно сворачиваться и принимать упорядоченную структуру. В конечном итоге белок высвобождается из рибосомы, свернутый в структуру, соответствующую его правильной конформации.

Разные члены семейства Hsp70 действуют на разные типы белков-мишеней. Цитоплазматические белки (собственно Hsp70 и родственный ему белок, получивший название Hsc70) действуют на насцентные белки на рибосомах. Варианты этих белков в ЭПР (называемые BiP или Grp78 у высших эукариот и Kar2 у *Saccharomyces cerevisiae*), а также шапероны митохондрий и хлоропластов похожим образом действуют на белки, проникающие посредством трансмембранного переноса внутрь органелл.

Какие особенности позволяют Hsp70 узнавать белок-мишень? Этот шаперон связывается с последовательностью аминокислот, заключенной в гидрофобный контекст. Поскольку в зрелом, приобретшем правильную конформацию белке такие мотивы запрятаны в глубь гидрофобного ядра, их экспозиция во внешнюю среду свидетельствует о том, что данный белок только что синтезирован либо денатурирован. Мотивы похожей структуры встречаются в среднем через каждые 40 аминокислот. Связывание предотвращает их неправильную агрегацию с другими участками белка.

Описанная модель взаимодействия шаперона с белком объясняет, как относящийся к семейству Hsp70 белок BiP может выполнять две функции: содействовать олигомеризации и/или фолдингу только что поступивших в ЭПР белков, а также удалять белки с неправильной структурой. Предположительно BiP распознает определенные участки полипептида, недоступные в зрелом, приобретшем правильную конформацию, белке. Эти участки экспонируются и привлекают BiP во время прохождения белка (в линейном виде) сквозь мембрану ЭПР. Если белок свернут неправильно или денатурирован, он может экспонировать на своей поверхности скрытые в норме последовательности.

10.7 Сигнальные последовательности инициируют транслокацию

Ключевые понятия:

- Белки попадают в систему ЭПР только котрансляционно.
- За ассоциацию с мембраной ЭПР ответственны сигнальные последовательности синтезируемых белков.

Белки, связывающиеся с мембранами посредством N-концевых лидерных последовательностей, используют широкий спектр сигналов, позволяющих им достичь своего конечного места назначения. В случае с ретикулоэндотелиальной системой окончательная локализация белка зависит от того, куда он будет направлен при прохождении



РИС. 10.14 Белки, поступившие в систему ЭПР–аппарат Гольджи могут быть транспортированы на плазматическую мембрану или, при наличии специфичных сигналов, к другому пункту назначения

через эндоплазматический ретикулум и аппарат Гольджи. С помощью лидерного пептида белок встраивается в мембрану и вследствие этого взаимодействия, проникает внутрь компартмента или на внутреннюю сторону мембраны (для мембрансвязанных белков). Мембрансвязанному белку необходим сигнал о прекращении его транспорта сквозь мембрану. Другие типы сигналов нужны для сортировки белков по различным компартментам или мембранам за пределами ЭПР. Процесс доставки белка к конечной точке его назначения, осуществляемый посредством транспорта через последовательные системы мембранных везикул, носит название **сортировки,** или **адресной доставки белков**.

На РИС. 10.14 изображена общая схема данного процесса. Секреторный путь включает в себя транспорт белка через ЭПР в аппарат Гольджи и далее на плазматическую мембрану. Белки, работающие внутри ЭПР, имеют С-концевой тетрапептид (KDEL). Этот пептид является сигналом для возвращения белка из аппарата Гольджи обратно в ЭПР. Сигналом, направляющим белок в лизосому, является ковалентное присоединение маннозо-6-фосфата. Чтобы остаться в аппарате Гольджи или попасть в структуру плазматической мембраны, белку требуются другие сигналы. Существует общая стартовая точка для белков, ассоциированных с или транспортируемых посредством мембран ретикулоэндотелиальной системы. *Эти белки связываются с мембраной ЭПР только непосредственно в процессе синтеза.* Рибосомы, синтезирующие мембрансвязанные белки, ассоциированны с мембраной ЭПР, что делает возможным котрансляционный перенос образующегося на них белка в мембрану. Участки ЭПР, с которыми связаны рибосомы, иногда называют «шероховатым ЭПР» (по внешнему виду этих мембран при электронной микроскопии), в противоположность «гладкому ЭПР», на поверхности которого расположено мало рибосом и который чаще имеет трубчатую, нежели пластинчатую, структуру. На РИС. 10.15 изображены мембрансвязанные рибосомы, синтезирующие белки внутрь ЭПР.

Мембранные белки, синтезируемые на шероховатом ЭПР, попадают из рибосомы непосредственно в мембрану ЭПР. Затем они в составе мембраны переносятся в аппарат Гольджи и оттуда направляются к своему конечному месту назначения: в лизосомы, секреторные везикулы

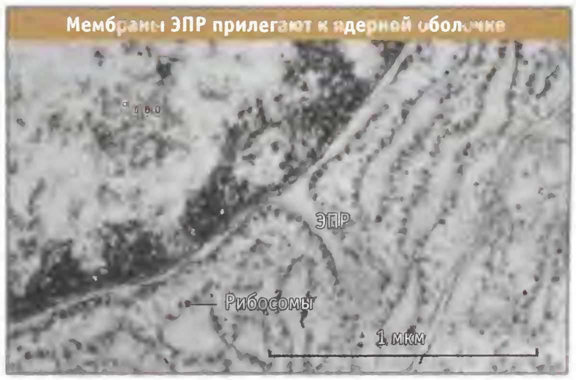


РИС. 10.15 Эндоплазматический ретикулум (ЭПР) представляет собой систему складчатых мембран, простирающихся в цитоплазму от ядра. Небольшие тельца, расположенные на внешней поверхности мембран, — рибосомы. Фотография любезно предоставлена Lelio Orci, University of Geneva, Switzerland



РИС. 10.16 Сигнальная последовательность бычьего гормона роста представлена 29 N-концевыми аминокислотами и состоит из центрального высоко гидрофобного региона, которому предшествуют (или окружают) регионы, состоящие из полярных аминокислот

или на плазматическую мембрану. Транспорт белков между органеллами идет при помощи мембранных везикул. Котрансляционная вставка белка в мембрану опосредована сигнальными последовательностями. Чаще всего это отщепляемая впоследствии N-концевая лидерная последовательность, состоящая из 15–30 аминокислот. На N-конце (или недалеко от него) расположены несколько полярных остатков, а внутри лидерной последовательности находится гидрофобное ядро, состоящее исключительно (или большей частью) из гидрофобных аминокислот. Каких-либо конкретных консервативных последовательностей для этого сигнала не выявлено. На РИС. 10.16 приведен пример такого сигнала.

Сигнальная последовательность необходима и достаточна для того, чтобы обеспечить перенос любого присоединенного к ней полипептида в компартмент-мишень. Сигнальная последовательность на N-конце белка глобина, к примеру, определяет тот факт, что он не остается в клетке, а транспортируется через клеточную мембрану.

Сигнальная последовательность определяет способность рибосомы связываться с мембраной ЭПР. Свободные рибосомы (синтезирующие белки в цитоплазме) и рибосомы, связанные с мембранами ЭПР, не отличаются друг от друга по своему строению. В момент начала синтеза белка на рибосоме, система еще не обладает информацией о том, останется ли синтезированный белок в цитоплазме или он будет синтезирован в ЭПР. Именно синтез сигнальной последовательности определяет связь рибосомы с мембраной.

10.8 Сигнальные последовательности взаимодействуют с сигнал-узнающими частицами (SRP)

Ключевые понятия:

- Сигнальные последовательности связываются с SRP (сигнал-узнающими частицами).
- Связывание SRP с сигнальной последовательностью белка вызывает остановку белкового синтеза.
- Синтез белка возобновляется, когда SRP связывается с SRP-рецептором на мембране ЭПР.
- Расположенная на внутренней поверхности мембраны сигнальная пептидаза отщепляет сигнальные последовательности от транспортируемого белка

Процесс транслокации белка можно разделить на две основные стадии: во-первых, рибосомы, несущие насцентный полипептид, связываются с мембранами ЭПР, а затем пептидная цепь перемещается в канал и транслоцируется сквозь него. Соединение рибосомы с мембраной требует участия **сигнал-узнающей частицы (SRP).** SRP обладает двумя важными особенностями:

- Она может связывать сигнальную последовательность растущего секреторного белка.
- Она может связывать расположенный на мембране ЭПР рецептор (SRP-рецептор).

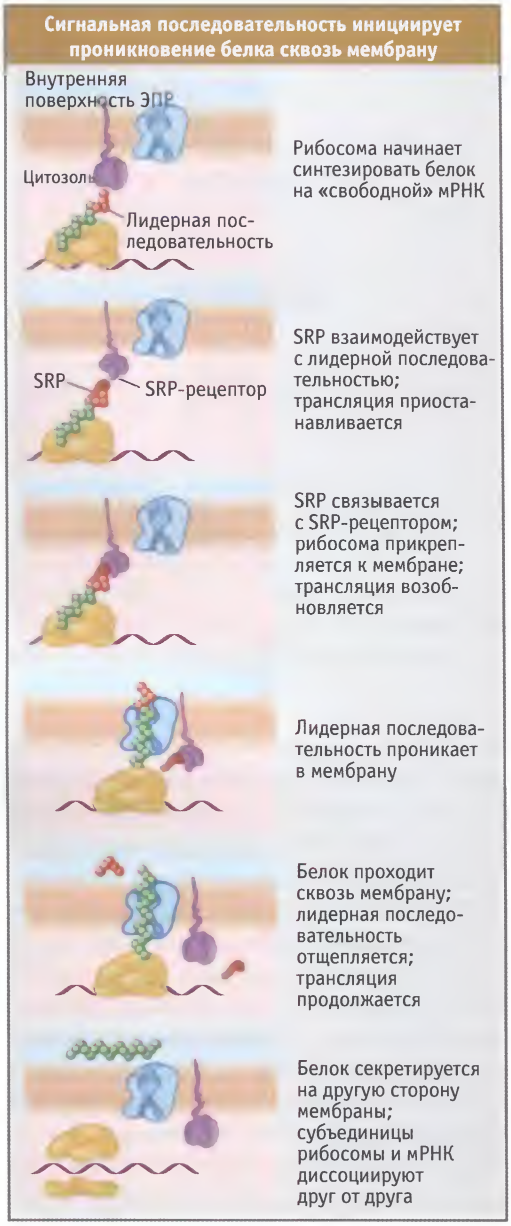


РИС. 10.17 Рибосомы, синтезирующие секреторные белки, ассоциируют с мембраной посредством сигнальной последовательности растущего полипептида

Функция SRP и SRP-рецептора заключается в доставке рибосом, несущих синтезируемый ими белок, к мембране. Первый шаг состоит в узнавании SRP сигнальной последовательности синтезируемого белка. Затем SRP взаимодействует с SRP-рецептором, и рибосома присоединяется к мембране. Все стадии трансляции мембранных белков приведены на РИС. 10.17.

SRP-рецептор выполняет промежуточную роль в транслокации белка. Связывание SRP с сигнальной последовательностью вызывает временную остановку трансляции. Как правило, это происходит после включения в полипептидную цепь ~70 аминокислот (в этот момент 25-аминокислотная лидерная последовательность уже

находится за пределами рибосомы, тогда как оставшиеся ~40 аминокислот еще скрыты внутри рибосомы).

После посадки SRP на SRP-рецептор сигнал-узнающая частица теряет сродство к сигнальной последовательности. При этом рибосома остается связана с ЭПР за счет других компонентов мембраны. Диссоциация SRP от рибосомы возобновляет процесс трансляции. После ассоциации рибосомы с мембраной SRP и SRP-рецептор уже сыграли свою роль, и оба фактора возвращаются в исходное состояние, снова приобретая способность связывать новый растущий полипептид с мембраной.

Временное прекращение синтеза белка (до момента ассоциации рибосомы с мембраной ЭПР) необходимо для контроля за конформацией белка. Если бы синтез секреторных белков происходил в цитоплазме, они могли бы принимать конформацию, не позволяющую им пройти сквозь мембрану. Способность SRP ингибировать трансляцию до тех пор, пока рибосома не свяжется с мембраной, позволяет предотвратить синтез белка, предназначенного для ЭПР, в цитоплазме.

Сигнальный пептид отщепляется от синтезируемого в ЭПР белка комплексом из пяти ферментов, называемых **сигнальными пептидазами.** Количество работающих в ЭПР комплексов сигнальных пептидаз в несколько раз превышает количество SRP и SRP-рецепторов. Их число примерно соответствует числу мембрансвязанных рибосом, из чего можно сделать вывод об ассоциации сигнальных пептидаз с трансмембранными синтезирующими комплексами. Пептидазы расположены на внутренней поверхности мембраны ЭПР и сигнальная последовательность должна целиком пройти сквозь мембрану, прежде чем подвергнется их действию. Гомологичные сигнальные пептидазы обнаружены у эубактерий, архей и эукариот.

10.9 SRP взаимодействует со своим рецептором

Ключевые понятия:

- SRP представляет собой комплекс из 7S РНК и шести белков.
- Бактериальным эквивалентом SRP является комплекс из 4,5S РНК и двух белков.
- SRP-рецептор представляет собой димер.
- После взаимодействия с рецептором SRP диссоциирует от него в результате гидролиза ГТФ.

Взаимодействие SRP с SRP-рецептором — ключевое событие в процессе ассоциации рибосомы (несущей насцентный пептид) с мембраной ЭПР в ходе трансляции у эукариот. Аналогичная система существуют и у бактерий, хотя ее роль в процессе более ограниченна.

SRP представляет собой 11S рибонуклеопротеиновый комплекс, состоящий из шести белков (общей массой 240 кДа) и малой (всего в 305 оснований, 100 кДа) 7S РНК. Как показано на **РИС. 10.18**, 7S РНК обеспечивает структурный каркас частицы, причем отдельные белки не ассоциируют друг с другом в ее отсутствие.

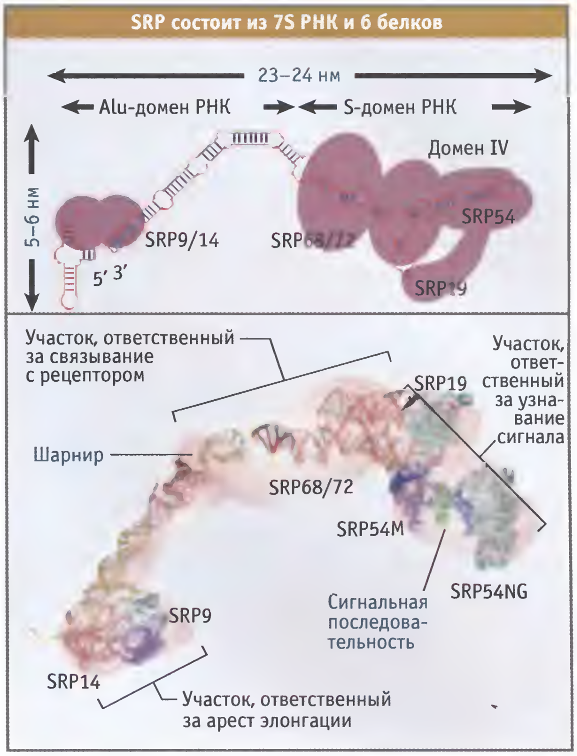


РИС. 10.18 7S РНК SRP имеет два домена. Белки связываются, как показано на диаграмме в верхней части рисунка. На нижней части рисунка приведена структура кристалла SRP. Каждая из функций SRP опосредована отдельным элементом структуры

В последовательности 7S РНК SRP частицы можно выделить два участка. 100 оснований на 5′-конце и 45 оснований на 3′-конце очень схожи с последовательностью Alu РНК (общей для млекопитающих малой РНК). В связи с этим их обозначают как **Alu-домен**. Оставшаяся часть 7S РНК SRP составляет **S-домен**.

Каждый из изображенных на рис. 10.18 элементов структуры SRP выполняет свою функцию в процессе остановки синтеза белка на рибосоме. SRP54 можно назвать наиболее важной субъединицей SRP. Она расположена на одном из концов молекулы 7S РНК и напрямую ответственна за узнавание субстратного белка путем связывания его сигнальной последовательности. Кроме того, совместно с димером SRP68-SRP72, расположенным в центральном регионе 7S РНК, SRP54 связывает SRP-рецептор. Димер SRP9-SRP14 расположен на другом конце молекулы и ответственен за остановку (арест) элонгации.

SRP является гибкой структурой. Изображенная на рис. 10.18 структура SRP, полученная на основе данных рентгеноструктурного анализа, демонстрирует, что в свободном состоянии (до ассоциации с сигнальной

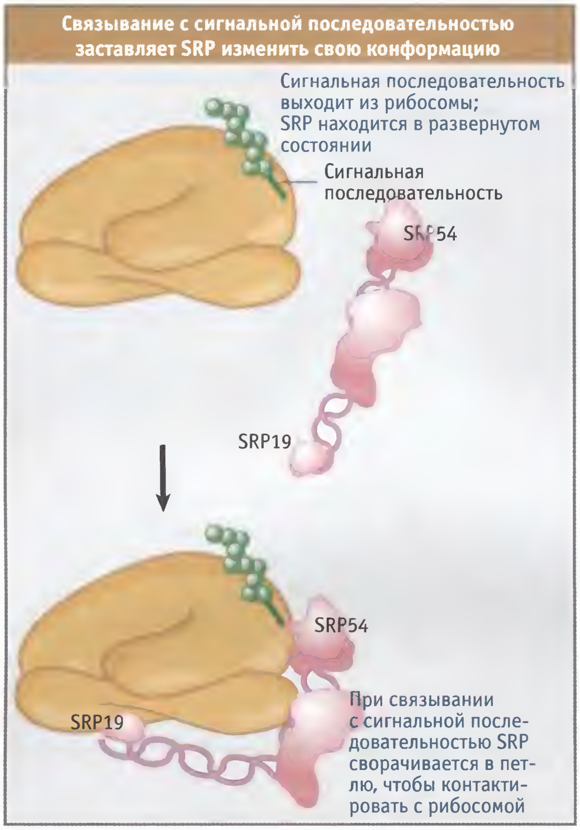


РИС. 10.19 SRP связывает сигнальную последовательность, как только она выходит из рибосомы. Это взаимодействие заставляет SRP сменить конформацию, изгибаясь на шарнире, что позволяет SRP54 контактировать с рибосомой в месте выхода белка, в это же время SRP19 формирует еще одну группу контактов с рибосомой

последовательностью) SRP имеет вытянутую форму. На РИС. 10.19 показано, как связывание с сигнальной последовательностью запускает процесс изменения конформации. Линейный комплекс сворачивается в петлю, что позволяет SRP54 контактировать с рибосомой в месте выхода синтезируемого белка, в то время как SRP19 располагается рядом с сайтом связывания фактора элонгации (определяя остановку элонгации и давая время для достижения сайта транслокации на мембране ЭПР).

SRP-рецептор представляет собой димер, состоящий из субъединиц SRα (72 кДа) и SRβ (30 кДа). β-субъединица заякорена в мембране ЭПР. N-конец большой α-субъединицы связан с β-субъединицей. Основная часть α-субъединицы выступает в цитоплазму. Бо́льшая часть последовательности цитоплазматического региона рецептора похожа на структуру белков, связывающих нуклеиновые кислоты (наличием положительно заряженных остатков). Данный факт позволяет предположить, что SRP-рецептор узнает в SRP именно 7S РНК.

Аналог SRP существует и у бактерий, однако он состоит из меньшего числа компонентов. У *E. coli* 4,5S РНК, взаимодействующая с рибосомами, аналогична 7S РНК SRP. Она связывается с двумя белками: Ffh (гомолог SRP54) и FtsY (гомолог α-субъединицы SRP-рецептора). В действительности FtsY одновременно выполняет функции и α-, и β-субъединицы; его N-концевой домен работает как SRPβ, ориентируя весь комплекс относительно мембраны, в то время как С-концевой домен взаимодействует с белком-мишенью. Роль этого комплекса несколько уже, чем роль комплекса SRP с SRP-рецептором. По всей видимости, комплекс позволяет части секреторных белков (но не всем) сохранять конформацию, что дает им возможность взаимодействовать с аппаратом секреции белков. Возможно, такая организация системы отражает первичную взаимосвязь между синтезом белка и его секрецией. У эукариот же SRP приобрела дополнительные функции задержки трансляции и посадки рибосомы на мембрану.

Почему же в составе SRP присутствует компонент РНК? Ответ должен крыться в эволюции SRP: ее происхождение относится к очень раннему этапу эволюции, к тому времени, когда доминировали ферментативные системы, основанные на РНК (аналогично рибосомам, чьи функции также опосредованы РНК). Кристаллическая структура комплекса между белок-связывающим доменом 4,5S РНК и РНК-связывающим доменом Ffh позволяет предположить, что РНК до сих пор продолжает играть важную роль в функционировании SRP.

4,5S РНК имеет в своем составе регион (домен IV), очень похожий на IV домен в 7S РНК (см. рис. 10.18). Белок Ffh состоит из трех доменов (N, G и M). Домен M (названный так вследствие высокого содержания метионина) выполняет ключевую функцию связывания с сигнальной последовательностью белка мишени (при помощи гидрофобного кармана). Гидрофобная сторона цепи метиониновых остатков формирует этот карман, взаимодействуя с бороздкой в структуре белка. За карманом следует мотив спираль-поворот-спираль, являющийся типичным для ДНК-связывающих белков (см. разд. 14.11 «Мотив «спираль-поворот-спираль» необходим репрессору для связывания ДНК»).

Рентгеноструктурный анализ позволил установить, как мотив спираль-петля-спираль домена M связывает дуплексный регион 4,5S РНК в IV домене. Отрицательно заряженный РНК-каркас примыкает к карману, повышая



РИС. 10.20 SRP, несущая ГДФ, связывает сигнальную последовательность. Затем рибосома стимулирует замещение ГДФ на ГТФ

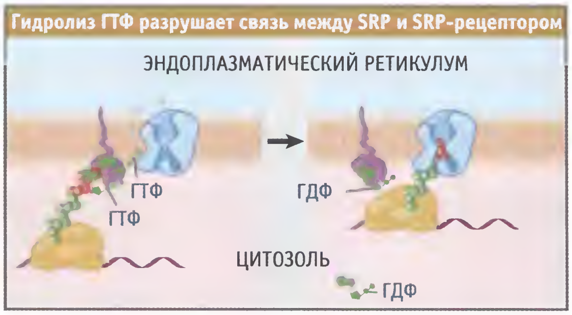


РИС. 10.21 И SRP, и SRP-рецептор проявляют ГТФазную активность для того, чтобы перенести сигнальную последовательность к мембране

вероятность того, что сигнальная последовательность действительно свяжется с обоими белками и РНК-компонентом SRP. Положительно заряженная последовательность, с которой начинается сигнальный пептид (см. рис. 10.16), может взаимодействовать с РНК, в то время как гидрофобный регион сигнального пептида будет оставаться в гидрофобном кармане.

Гидролиз ГТФ играет важную роль в процессе посадки рибосомы (с сигнальной последовательностью) на мембрану ЭПР. И SRP, и SRP-рецептор обладают ГТФазной активностью. С структуре SRP ГТФазой является сигнал-связывающая субъединица SRP54. ГТФазами являются и обе субъединицы SRP-рецептора. ГТФазная активность этих участников процесса необходима для того, чтобы рибосома связалась с мембраной ЭПР. На РИС. 10.20 показано, как SRP в комплексе с ГДФ связывает сигнальную последовательность, после чего рибосома стимулирует замену ГДФ на ГТФ. Причем сигнальная последовательность синтезируемого белка ингибирует гидролиз новой молекулы ГТФ, в результате чего комплекс остается ГТФ-связанным до момента встречи с SRP-рецептором.

Для того чтобы растущий белок был перенесен от SRP на мембрану, SRP должна потерять связь с SRP-рецептором. Как показано на РИС. 10.21, для осуществления этого процесса необходимо, чтобы ГТФазную активность проявляли как SRP, так и SRP-рецептор. Данная реакция была описана для бактериальной системы и отличается следующей особенностью: белок Fts активировал гидролиз ГТФ белка Fth, а белок Fth реципрокно активирует гидролиз белка FfsY.

10.10 Транслокон формирует пору

Ключевые понятия:

- Тримерный комплекс Sec61 формирует канал для транспорта белков через мембрану.
- Транспортируемый через мембрану белок поступает в транслокон непосредственно с рибосомы (без контакта с цитоплазмой).

Перенос (как правило) гидрофильных белков через гидрофобную мембрану связан со следующей проблемой: взаимодействие между заряженным белком и гидрофобными липидами чрезвычайно невыгодно с энергетической точки зрения. Было установлено, что транслоцируемый через мембрану ЭПР белок можно выделить с помощью денатурирующих агентов (эффективно действующих в водном растворе). Однако с помощью тех же детергентов не удается выделить белки, являющиеся постоянными компонентами мембраны. На основе этих данных была предложена модель транслокации, изображенная на РИС. 10.22, на котором показано, что мембранные белки ЭПР формируют гидрофильный канал, пронизывающий липидный бислой. Транспортируемый белок перемещается по этому каналу, взаимодействуя только с формирующими канал гидрофильными аминокислотами, но не с липидами. С внутренней стороны (в просвет цистерны ЭПР) мембраны канал закрыт, что предотвращает свободное перемещение ионов между ЭПР и цитоплазмой.

Пронизывающий мембрану канал получил название **транслокон.** Его компоненты были установлены при помощи двух подходов: 1) метода белковых сшивок, который позволил выделить белки мембраны ЭПР, образующие ковалентные сшивки (в определенных условиях) с транспортируемыми белками, потенциально являющимися субъединицами канала; и 2) метода анализа мутантов у дрожжей, который позволил обнаружить мутантов sec (названных так вследствие нарушенной секреции белков), содержащих в своем составе класс белков, вызывающих накопление предшественников секреторных и мембранных белков в цитоплазме. С помощью этих двух подходов был идентифицирован комплекс Sec61, состоящий из трех трансмембранных белков: Sec61α, β и γ. Sec61 является основным компонентом транслокона. В детергенте (обеспечивающем гидрофобную среду, имитирующую условия мембранного окружения) Sec61 формирует цилиндрические

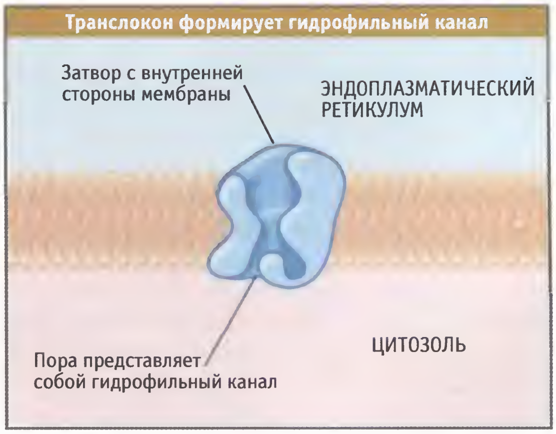


РИС. 10.22 Транслокон — тримерный Sec61 комплекс, формирующий мембранный канал. Канал закрыт с внутренней стороны мембраны

олигомеры с диаметром ~85 Å и центральной порой диаметром ~20 Å, причем каждый олигомер состоит из четырех гетеротримеров.

Похожей тримерной структурой обладают мембранные каналы всех групп организмов. В бактериях и археях такую структуру назвали SecY-комплексом. Субъединица Sec61α у эукариот (как и соответствующая ей у бактерий и архей субъединица SecY) имеет наиболее консервативную последовательность и обеспечивает формирование поры, сквозь которую транспортируется белок. Пора формируется из димеров тримерного комплекса, ориентированных друг относительно друга «спина-к-спине» таким образом, что α-субъединицы формируют единый канал. Белковый комплекс, формирующий канал в мембране, может состоять и из двух димеров, ориентированных друг относительно друга «бок-о-бок», однако в этом случае только один из них доступен для взаимодействия с рибосомой.

Имеет ли канал заранее определенную структуру (как это изображено на рисунке) или же она формируется в ответ на взаимодействие гидрофобной сигнальной последовательности с липидным бислоем? Каналы могут быть обнаружены благодаря их способности пропускать ионы (оцениваемой как локальное изменение электрической проводимости). Ионопроводящие каналы могут быть детектированы в мембране ЭПР, и их состояние зависит от транслокации белков. Данный факт доказывает, что каналы являются постоянным компонентом мембраны.

В момент переноса растущего полипептида от рибосомы к мембране ЭПР канал открывается. Транспортируемый белок полностью заполняет канал таким образом, что ионы не могут проникать сквозь канал во время транслокации. Однако если белок высвобождается из канала (например, в результате действия пуромицина), канал становится проницаем для ионов. Когда рибосома диссоциирует с мембраны, канал закрывается, что указывает на участие рибосомы в поддержании канала в открытом состоянии. Скорее всего, состояние канала в данном случае зависит от наличия транспортируемого белка.

Оценка способности флуоресцентно меченных агентов разных размеров проходить сквозь канал демонстрирует, что сам канал имеет довольно большой внутренний диаметр (в диапазоне от 40 до 60 Å). Это намного больше, чем диаметр α-спиральной структуры белка. Более того, эти значения превышают видимый размер поры при исследовании методами электронной микроскопии. Данное противоречие пока остается неразрешенным.

Окружение аминокислоты в белке можно оценить при помощи включения данной аминокислоты с фотореактивными остатками. По флуоресценции этих остатков можно судить, находятся они в гидрофильной или же в гидрофобной среде. Эксперименты с такими пробами показывают, что в процессе синтеза на рибосоме сигнальная последовательность находится в гидрофильной среде, однако недоступна для ионов цитоплазмы. Она остается в гидрофильной среде вследствие своего взаимодействия с каналом на мембране. Это позволяет предположить, что в ходе синтеза белок переносится из замкнутого рибосомального туннеля непосредственно в гидрофильную среду мембранного канала.

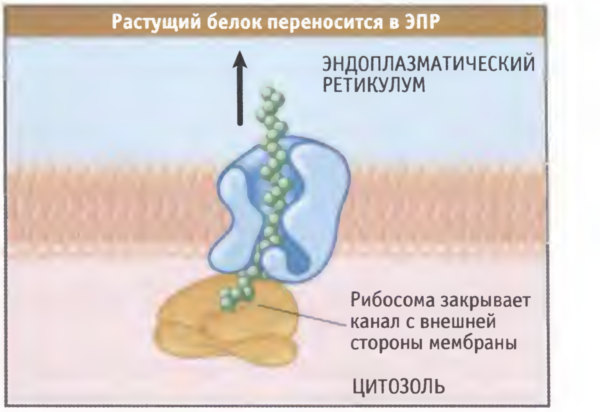


РИС. 10.23 Растущий (насцентный) белок переносится к транслокону прямо от рибосомы. Рибосома закрывает канал с внешней стороны мембраны

В действительности доступ к поре (или «пропускание») контролируется с *обеих* сторон мембраны. До прикрепления рибосомы пора закрыта с внутренней стороны ЭПР. На РИС. 10.23 изображено, как прикрепление рибосомы обеспечивает закрывание поры со стороны цитоплазмы. Когда длина растущего белка достигает ~ 70 аминокислот (что соответствует протягиванию полипептидной цепи через весь канал), пора открывается с внутренней стороны. Таким образом, в любой момент времени пора закрыта с одной из сторон, что способствует поддержанию постоянства ионного состава в каждом из компартментов.

Транслокон является мультифункциональной структурой и может использоваться транслоцируемым белком в нескольких целях:

- Как средство переноса синтезируемых белков с цитоплазматических рибосом в просвет ЭПР (см. разд. 10.11 «Для транслокации необходимы дополнительные включения в транслокон, а иногда и храповик ЭПР»).
- С помощью транслокона интегральные мембранные белки системы ЭПР встраиваются в мембрану; для этого необходимо, чтобы канал был открыт или дезагрегирован таким образом, чтобы белок в липидном бислое мог двигаться латерально (см. разд. 10.15 «Как осуществляется вставка белков в мембрану?»).
- В ряде случаев белки могут переноситься из просвета ЭПР обратно в цитоплазму; этот процесс известен как «обратная транслокация» (см. разд. 10.12, «Для деградации белки возвращаются в цитозоль путем обратного транспорта»).

10.11 Для транслокации необходимы дополнительные включения в транслокон, а иногда и храповик ЭПР

Ключевые понятия:

- Рибосома, SRP и SRP-рецептор создают условия для попадания насцентного белка в транслокон.
- Белки, попадающие в транслокон посттранляционно, требуют наличия дополнительных компонентов в цитоплазме и белка BiP в ЭПР.
- BiP — это храповик, предотвращающий движение белка в обратную сторону.

Транслокон и SRP-рецептор являются основными элементами котрансляционной транслокации. При встраивании комплекса Sec61 вместе с SRP-рецептором в искусственную мембрану, он оказывается способен поддерживать транслокацию некоторых насцентных белков. Однако большинство синтезируемых белков требуют присутствия дополнительного компонента — белка, связывающего траслоцируемую насцентную цепь с мембраной (белок TRAM). В экспериментах этот белок чаще всего оказывается ковалентно пришит к транспортируемой через мембрану растущей полипептидной цепи. Установлено, что TRAM усиливает транслокацию любых белков.

Все компоненты транслокона и выполняемые ими функции суммированы на РИС. 10.24. Простота искусственной системы позволяет сделать несколько важных выводов. Мы представили Sec61 как комплекс, формирующий канал и взаимодействующий с рибосомой. Процесс начинается с узнавания SRP сигнальной последовательности в момент выхода растущей полипептидной цепи из рибосомы. Далее SRP связывается с SRP-рецептором, и сигнальная последовательность переносится в транслокон. Когда сигнальная последовательность поступает внутрь транслокона, рибосома присоединяется к комплексу Sec61, закрывая собой конец поры, обращенный в цитоплазму. Поскольку в искусственной системе отщепления сигнального пептида не происходит, можно сделать вывод о необязательности данного этапа для процесса транслокации. Также при помощи системы *in vitro* показано, что компоненты внутренней стороны мембраны не участвуют в транслокации.

Конечно, эффективность работы системы *in vitro* довольно низка. В системе *in vivo* масса дополнительных компонентов участвует в повышении эффективности процесса, а также предотвращает вмешательство в процесс других клеточных белков.

В случаях, когда транспорт белка через мембрану происходит посттрансляционно, требуется участие более сложного аппарата. Тот же Sec61-комплекс формирует канал, однако кроме него необходимы еще четыре дополнительных белка Sec. Вместе с тем на внутренней стороне мембраны должен присутствовать шаперон BiP (белок связывания, член семейства Hsp70) и АТФ. На РИС. 10.25 показано, что BiP функционирует в качестве своеобразного храповика. В отсутствие BiP транслоцируемый белок вследствие броуновского движения соскальзывает обратно в цитоплазму. Располагаясь на внутренней стороне мембраны ЭПР, BiP захватывает белок, как только тот выходит из поры в просвет ЭПР. Это предотвращает движение белка в обратном направлении. BiP не тянет белок сквозь мембрану, он лишь не дает ему скользить обратно. (Причина, по которой BiP требуется для посттрансляционной транслокации, но не является необходимым для осуществления котрансляционной транслокации, может заключаться в том, что синтезируемый белок, постоянно выталкиваясь из рибосомы, не может двигаться в обратном направлении.)

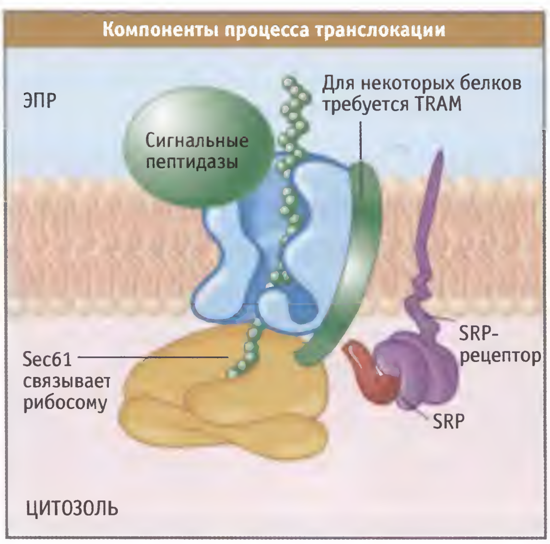


РИС. 10.24 В процессе транслокации участвуют транслокон, SRP, SRP-рецептор, Sec61, TRAM и сигнальные пептидазы

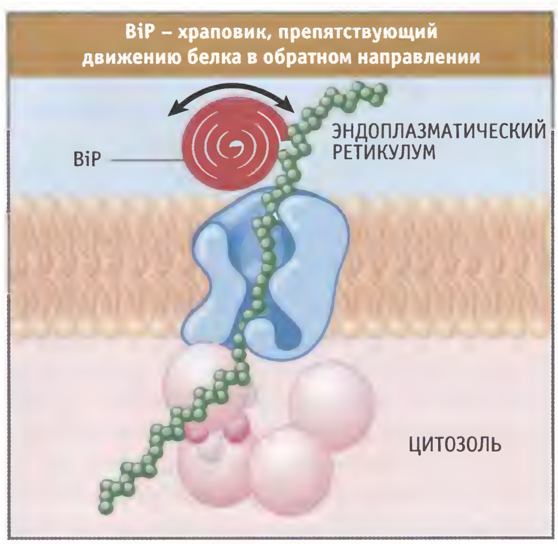


РИС. 10.25 Белок BiP действует в качестве храпового механизма, предотвращая обратную диффузию переносимого белка

10.12 Для деградации белки возвращаются в цитозоль путем обратного транспорта

Ключевые понятия:

- Транслокон Sec61 может использоваться для обратного переноса белков из ЭПР в цитоплазму.

Внутри ЭПР происходит несколько важных процессов: белки движутся по ЭПР в разных направлениях, они гликозилируются и принимают свои конечные конформации; параллельно ЭПР обеспечивает систему «контроля качества», в которой неправильно свернувшиеся белки распознаются и подвергаются деградации. Однако собственно деградация белков происходит и вне ЭПР, для ее реализации необходим экспорт белка обратно в цитоплазму.

Первое указание на то, что белки деградируют в цитоплазме, а не в ЭПР, было получено на основании факта вовлечения в процесс деградации протеасом (больших белковых агрегатов с несколькими протеолитическими активностями). Ингибиторы протеасом препятствуют деградации бракованных белков ЭПР. Сигналом для расщепления белка протеасомой служит присоединение к нему убиквитина — небольшой полипептидной цепи. Важно отметить, что убиквитинирование и протеасомная деградация происходят в цитоплазме (и частично в ядре).

Транспорт белка из ЭПР в цитоплазму является процессом, обратным импорту белка в ЭПР. Он получил название **обратной транслокации** (иногда его называют ретротранслокацией или дислокацией). Транслокон Sec61 принимает участие в обратной транслокации. Однако условия протекания процессов прямой и обратной транслокации различны; к примеру при обратной транслокации транслокон не ассоциирует с рибосомой. Некоторые мутации в Sec61 нарушают обратную транслокацию, но не влияют на прямую транслокацию. Это можно объяснить различиями в механизмах прямой и обратной транслокаций либо (что более вероятно) тем, что регионы, затрагиваемые этими мутациями, отвечают за взаимодействие с дополнительными компонентами системы обратной транслокации.

До сих пор неясно, каким образом канал открывается и пропускает белок с внутренней стороны ЭПР (см. РИС. 10.26). Предполагается, что в этом процессе участвуют специальные компоненты обратной транслокации. Согласно одной из моделей, неправильно свернувшиеся или неправильно синтезированные белки узнаются шаперонами, которые и переносят их к транслокону. Так, цитомегаловирус человека (ЦМВ) кодирует цитозольные белки, разрушающие синтезирующиеся молекулы белка МНС (главного комплекса гистосовместимости) I класса. Этот процесс требует от вируса продукции белка US11, который, являясь мембранным белком, функционирует в ЭПР. Взаимодействуя с молекулами МНС этот белок, возможно, направляет их в транслокон для обратной транслокации.



РИС. 10.26 В результате обратной транслокации несвернутый белок по транслокону перемещается из ЭПР в цитозоль для деградации. Механизм вставки белка в транслокон не известен

Систему деградации белков у дрожжей можно исследовать с помощью мутаций, приводящих к накоплению абберативных белков. В большинстве случаев неверно свернувшийся белок не транспортируется через систему ЭПР–комплекс Гольджи, а деградирует. Мутантные дрожжи, неспособные деградировать бракованные белки, можно разделить на два класса: одни демонстрируют нарушения компонентов протеолитического аппарата (таких как ферменты, участвующие в убиквитинировании); другие — в компонентах транспортного аппарата (включающих Sec61, BiP и Sec63).

Белки, вовлеченные в процесс деградации, можно выявить на основании их взаимодействия с белками цитомегаловируса. Вирусный белок US11 переносит МНС-субстрат к белку (называемому Derlin), находящемуся на мембране ЭПР. Derlin, в свою очередь, транспортирует субстрат к цитозольной АТФазе — p97, которая, по всей видимости, ответственна за вытягивание белка из канала в цитоплазму. Derlin является гомологом дрожжевого белка Der1p (одного из белков определенного методом изучения мутантных клеток как часть системы обратной транслокации).

10.13 Белки заякорены в мембране при помощи гидрофобных регионов

Ключевые понятия:

- У мембранных белков I группы N-конец молекулы расположен на внутренней стороне мембраны, а белки II группы имеют противоположную ориентацию.
- Некоторые белки имеют несколько мембрансвязанных доменов.

Все биологические мембраны содержат белки, удерживаемые в липидном бислое посредством нековалентных взаимодействий. Согласно принятому определению **интегральных мембранных белков**, они могут быть из-

влечены из мембраны только при разрушении липидного бислоя. Общей чертой таких белков является наличие хотя бы одного **трансмембранного домена,** представленного α-спиральным участком длиной в 21–26 гидрофобную аминокислоту. Последовательность аминокислот, удовлетворяющая критериям встраивания в мембрану, может быть определена при помощи диаграммы гидрофобности, позволяющей оценить суммарную гидрофобность участка полипептида. Белок, имеющий участки, выступающие на поверхность с обеих сторон мембраны, называется **трансмембранным белком**. Ассоциация белка с мембраной может принимать различную форму. Топография мембранного белка зависит от числа и конфигурации трансмембранных регионов.

Если белок имеет единственный трансмембранный регион, его позиция определяется тем, насколько белок выступает наружу с каждой из сторон мембраны. Белок может иметь крупные домены, выступающие с обеих сторон мембраны, либо, если трансмембранный регион расположен близко к концу молекулы, выступающий с одной из сторон мембраны домен будет небольшим или вовсе не будет экспонирован. Длина N-концевого или C-концевого участка, выдающегося за пределы мембраны рядом с трансмембранным регионом, варьирует в широких пределах.

Как показано на РИС. 10.27, белки, имеющие единственный трансмембранный домен, делят на две группы. Белки I группы, у которых во внеклеточное пространство обращен N-концевой участок молекулы, встречаются чаще, чем белки II группы, имеющие обратную ориентацию (с N-концом, обращенным в цитоплазму). Ориентация белковой молекулы определяется в момент прохождения белка сквозь мембрану ЭПР.

На РИС. 10.28 показано, какие ориентации принимают белки, имеющие несколько трансмембранных доменов. Нечетное число таких доменов означает, что концы молекулы белка окажутся на противоположных сторонах мембраны, четное число доменов предполагает, что концевые участки будут располагаться с одной стороны. Протяженность участков, выступающих с одной или с обеих сторон мембраны, определяется местоположением трансмембранных доменов. Домены по концам молекулы могут выступать далеко за пределы мембраны, а последовательности, расположенные между доменами, — выпячиваться в водную фазу в виде петель. Одним из наиболее распространенных типов структуры является трансмембранный пассаж из семи доменов (или «серпантиновый» рецептор); другой пример — трансмембранный пассаж из 12 доменов белка — компонента ионного канала.

Играет ли трансмембранный домен еще какую-нибудь роль в функционировании белка кроме обеспечения его встраивания в липидный бислой? В простых белках I и II групп он либо не играет никакой дополнительной роли, либо выполняет незначительную функцию; часто он может быть заменен любым другим трансмембранным доменом. Однако в белках, многократно пересекающих мембрану либо имеющих субъединицы, олигомеризующиеся внутри мембраны, трансмембранные домены играют важную роль. В этих случаях они часто содержат полярные

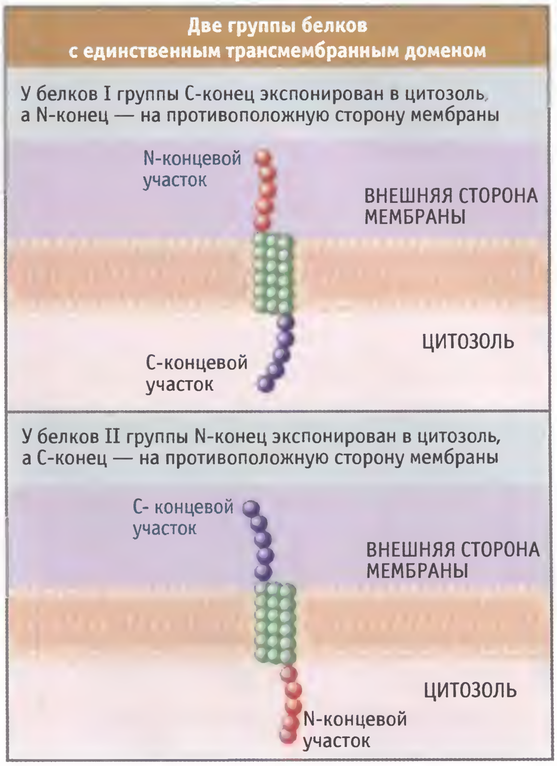


РИС. 10.27 Трансмембранные белки I и II групп имеют противоположную ориентацию по отношению к мембране

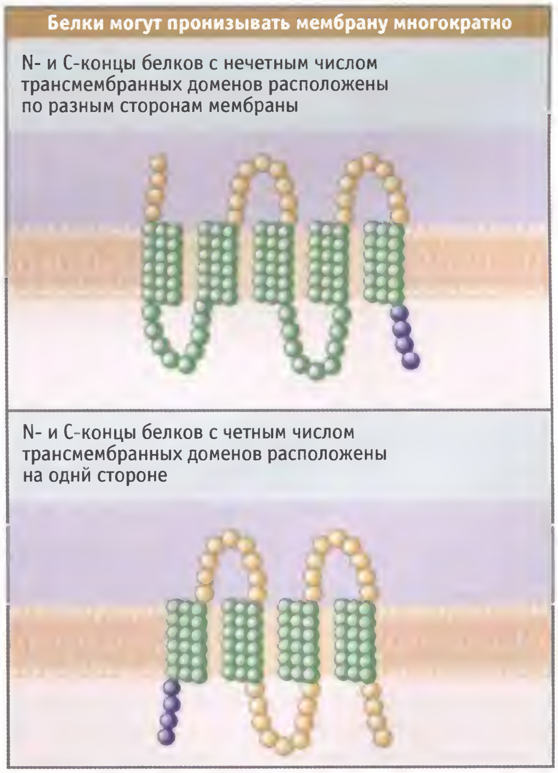


РИС. 10.28 Ориентация концов молекулы белков, неоднократно пронизывающих мембрану, зависит от количества трансмембранных сегментов

остатки, не обнаруживаемые в одиночных трансмембранных доменах белков I и II групп. Полярные регионы в мембраносвязанных доменах не взаимодействуют с липидным бислоем, вместо этого они взаимодействуют друг с другом. Это дает белкам возможность формировать пору или канал внутри липидного бислоя. Благодаря взаимодействию между такими доменами образуется гидрофильный канал через гидрофобную среду мембраны, что, в свою очередь, делает возможным прохождение через мембрану заряженных ионов и молекул и имеет большое значение для функционирования ионных каналов и транспорта разнообразных лигандов. Другим примером, демонстрирующим роль конформации трансмембранных доменов, являются рецепторы, связывающие липофильные лиганды. В этих случаях сами трансмембранные домены (а не внеклеточные домены) связывают лиганды непосредственно внутри структуры мембраны.

10.14 Якорные последовательности определяют ориентацию белка в мембране

Ключевые понятия:

- Якорная последовательность останавливает движение белка через транслокон. Как правило, она расположена на С-конце молекулы, приводя к образованию мембранного белка I группы с N-концом, прошедшим сквозь мембрану и экспонированным на внутренней ее стороне.
- Комбинированная сигнально-якорная последовательность служит одновременно и для введения белка сквозь мембранный канал, и для заякоривания в мембране вошедшего участка молекулы. Поскольку заякоривание идет с внутренней стороны мембраны, мембранный белок приобретает ориентацию белка II группы с N-концом, экспонированным в цитоплазму.

Белки, секретируемые клеткой, проходят сквозь мембрану благодаря гидрофильной среде канала транслокона. В противоположность им, мембраносвязанные белки, начиная процесс таким же образом, затем переносятся из гидрофильной среды канала в гидрофобную среду окружающей мембраны. Основным вопросом, возникающим при рассмотрении механизма встраивания белков в мембрану, является объяснение причин такого переноса на основе различий между трансмембранными и секреторными белками. Путь, посредством которого белки как I, так и II группы встраиваются в мембрану, напоминает механизм движения через мембрану секреторных белков, основывающийся на котрансляционном функционировании сигнальной последовательности. В отличие от секреторных белков, белки, остающиеся в мембране, обладают дополнительным **стоп-транспортным** сигналом. Он представлен группой гидрофобных аминокислот, примыкающих к нескольким заряженным аминокислотам. Эта последовательность выполняет роль **якоря,** цепляясь за мембрану и не давая белку уйти от нее.

Удивительное свойство якорных последовательностей заключается в том, что, в зависимости от местоположения в структуре полипептида, они могут функционировать и как сигнальные последовательности. В случае когда якорная последовательность находится в составе белка, дефицитного по другим типам сигналов, она принимает на себя и функцию обеспечения транслокации. Возможное объяснение такой двойной функции якорных последовательностей заключается в том, что сигнальная и якорная последовательности взаимодействуют с некоторыми общими компонентами аппарата транлокации. Связывание этих компонентов с сигнальной последовательностью инициирует транслокацию, а затем якорная последовательность вытесняет сигнальную и останавливает перенос белка.

Интеграция белка в мембрану начинается со встраивания сигнальной последовательности (имеющей форму шпильки), при этом N-конец белка остается на цитоплазматической стороне мембраны. Позицию и ориентацию белка в мембране определяют два параметра: отщепляется ли сигнальная последова-

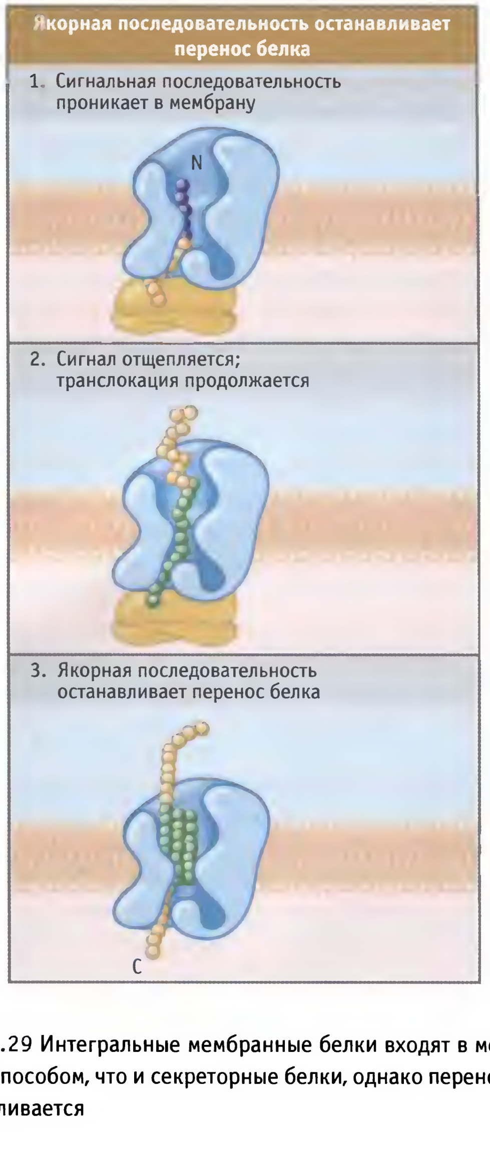


РИС. 10.29 Интегральные мембранные белки входят в мембрану тем же способом, что и секреторные белки, однако перенос белка останавливается

тельность и каково местоположение якорной последовательности?

Процесс встраивания белков I группы проиллюстрирован на РИС. 10.29. Сигнальная последовательность расположена на N-конце молекулы, а местоположение якорного сигнала определяет момент остановки транслокации белка. После интеграции якорной последовательности в мембрану N-концевые домены молекулы белка будут обращены внутрь органеллы, в то время как домены С-конца — в цитоплазму.

Как правило, стоп-сигналы такого типа расположены на С-конце. Как видно из рисунка, транслокация в данном случае останавливается только после прохождения через мембрану последних участков молекулы. Такая схема характеризует расположение в мембране многих белков, в том числе и белков поверхности клетки. Большая часть последовательности такого белка экспонирована на внутренней поверхности мембраны ЭПР, и лишь совсем небольшая часть обращена в цитоплазму.

Белки II группы не имеют на N-конце отщепляемой (впоследствии) лидерной последовательности. Вместо этого сигнальная последовательность комбинирована с якорной последовательностью. Обобщенный механизм интеграции белков II типа в мембрану включает в себя стадии, изображенные на РИС. 10.30. В противоположность проходящей сквозь мембрану сигнальной последовательности, комбинированная сигнально-якорная последовательность не выходит за пределы мембраны. Вместо этого она остается в мембране (возможно, напрямую взаимодействуя с липидным бислоем), в то время как оставшаяся часть растущего полипептида продолжает выпячиваться в просвет ЭПР.

Сигнально-якорная последовательность, как правило, является внутренней, и ее местоположение определяет, какая часть белка останется в цитозоле, а какая окажется во внеклеточном пространстве. Все N-конце-

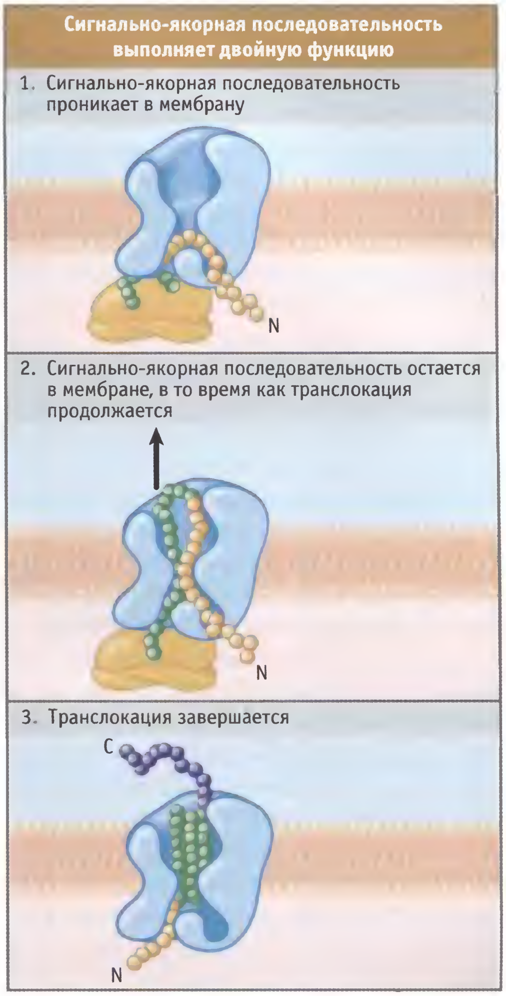


РИС. 10.30 Комбинированная сигнально-якорная последовательность заставляет белок менять свою ориентацию таким образом, что N-конец молекулы остается на внутренней поверхности мембраны, а С-конец — на внешней



РИС. 10.31 Сигнально-якорная последовательность нейраминидазы вируса гриппа расположена близко к N-концу и имеет гидрофобное ядро

вые последовательности, предшествующие сигнально-якорной последовательности, будут экспонированы в цитоплазму. Чаще всего цитоплазматический участок белковой молекулы сравнительно короток — от 6 до 30 аминокислот. Поскольку N-конец молекулы остается в связанном с мембраной состоянии, пока оставшаяся часть белковой молекулы проходит сквозь мембрану, это меняет ориентацию белка по отношению к мембране.

Комбинированная сигнально-якорная последовательность белков II группы напоминает отщепляемую сигнальную последовательность белков I группы. На РИС. 10.31 приведен пример сигнально-якорной последовательности. Как и в случае с отщепляемой лидерной последовательностью, природа аминокислот, входящих в состав сигнально-якорной последовательности, имеет большее значение, чем фактическая последовательность. Регионы, расположенные на концах сигнально-якорной последовательности, несут положительный заряд; центральные регионы не заряжены и напоминают гидрофобное ядро отщепляемого лидерного пептида. Мутации, привносящие заряженные аминокислоты в центральный регион сигнально-якорной последовательности, препятствуют ее встраиванию в мембрану; мутации на любом из концов также мешают процессу заякоривания, вследствие чего белок либо секретируется, либо накапливается в ошибочном компартменте клетки.

Заряд аминокислот, окружающих якорную последовательность, оказывает существенное влияние на ориентацию белка. Как правило, домен с бóльшим положительным зарядом оказывается на цитоплазматической стороне мембраны (сторона, где расположен N-конец белков II группы). Если в результате мутаций положительные заряды пропадают, ориентация белка может поменяться на противоположную. Влияние заряда на ориентацию белковой молекулы описывается правилом «положительное внутри» (positive inside), которое гласит, что конец якорной последовательности с наибольшим положительным зарядом будет расположены в цитоплазме. Положительно заряженные последовательности в действительности являются средством закрепления на цитоплазматической стороне мембраны, что, в свою очередь, контролирует направление вставки в мембрану гидрофобного региона и общую ориентацию белка.

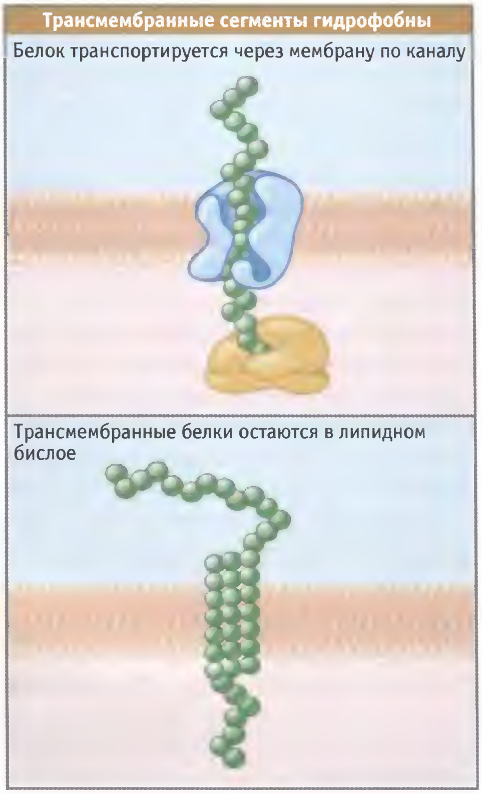


РИС. 10.32 Как из белкового канала трансмембранный белок переходит в липидный бислой?

10.15 Как белки встраиваются в мембраны?

Ключевые понятия:

- Переход трасмембранных доменов из транслокона в липидный бислой запускается в результате взаимодействия трансмембранного региона с транслоконом.

К настоящему моменту исследователи пришли к пониманию механизмов, при помощи которых секреторные белки проходят сквозь мембрану, а также того, как эти механизмы соотносятся с процессами встраивания в мембрану белков I и II групп, обладающих единственным трансмембранным доменом. Однако до сих пор не установлены детали процесса интеграции в мембрану белков с несколькими трансмембранными доменами.

Мы хорошо представляем, как секреторные белки проходят сквозь мембрану, однако применить эту же модель к белкам, остающимся в мембране, затруднительно. РИС. 10.32 демонстрирует различия между организацией транспортируемых белков (защищенных от контакта с липидным бислоем посредством гидрофильного канала) и трансмембранных белков, обладающих гидрофобными сегментами, напрямую взаимодействующими с мембраной. Принципиальным остается вопрос о механизме переноса белка из гидрофильного канала в липидный бислой.

Возможный механизм переноса гидрофобных трансмембранных доменов напрямую из канала в мембрану изображен на РИС. 10.33. Эта идея согласуется с наблюдениями в системах, с помощью которых оценивался перенос в липидное окружение белков с разными трансмембранными доменами *in vitro*. Когда домен проходит порог гидрофобности, белок из канала, состоящего из комплекса Sec61 и TRAM, может перейти в липидный бислой. Кроме общей гидрофобности важный эффект оказывает и расположение полярных остатков внутри трансмембранных сегментов. Самое простое объяснение заключается в том, что структура канала



РИС. 10.33 Свежесинтезированные мембранные белки способны перемещаться латерально из транслокона в липидный бислой. Механизм этого переноса не известен

позволяет транслоцируемому белку напрямую контактировать с липидным бислоем, так, что соответствующие гидрофобные сегменты могут сразу переходить в липидное окружение. Структура транслокона предполагает существование «ворот», расположенных между двумя спиралями и используемых для переноса участка полипептидной цепи в бислой.

До последнего времени исследователи считали, что независимо от того, как именно происходит перенос трансмембранного сегмента в мембрану, этот процесс запускается самим транслоцируемым пептидом, находящимся в поре. Однако оказалось, что изменения в структуре поры появляются раньше, еще в момент синтеза трансмембранной последовательности на рибосоме. Как уже было сказано, пока секреторный белок проходит сквозь пору, с цитозоплазматической стороны мембраны канал закрыт рибосомой на протяжении всего синтеза, однако с внутренней стороны мембраны канал открывается после синтеза первых семидесяти аминокислотных остатков. При синтезе трансмембранных белков, как только трансмембранный регион полностью синтезирован (несмотря на то что он еще целиком находится внутри рибосомы), пора закрывается с внутренней стороны. Как это изменение связано с переносом трансмембранной последовательности в мембрану, не установлено.

Процесс интеграции в мембрану хорошо описан для белков I и II групп, имеющих единственный трансмембранный домен. Однако о встраивании в мембрану белков с несколькими мембран-связывающими доменами известно гораздо меньше. Считают, что процесс также опосредован последовательностями с сигнальными и/или якорными свойствами. Согласно одной из моделей, в полипептиде присутствуют серии перемежающихся сигнальных и якорных последовательностей. Транслокация инициируется первой сигнальной последовательностью и продолжается до первой якорной последовательности. Затем процесс реинициируется следующей сигнальной последовательностью и снова продолжается до очередной якорной последовательности. В некоторых случаях трансмембранный домен перемещается в липидный бислой сразу же после попадания в транслокон, что подразумевает существование разных путей интеграции белка в мембрану.

10.16 Посттрансляционная интеграция белков в мембрану зависит от лидерных последовательностей

Ключевые понятия:

- N-концевая лидерная последовательность несет информацию, необходимую для ассоциации белка с мембранами митохондрий или хлоропластов.

Митохондрии и хлоропласты самостоятельно синтезируют только некоторые из своих белков. Так, в митохондриях синтезируется лишь около 10 белков органеллы; в хлоропластах — порядка 50. Большинство же белков этих органелл синтезируется в цитоплазме тем же пулом свободных рибосом, которые синтезируют все цитоплазматические белки. Следовательно, после синтеза в цитоплазме эти белки должны быть транспортированы в органеллы.

Большинство белков, попадающих в митохондрии и хлоропласты посттрансляционно, имеют лидерные последовательности, отвечающие за первичное связывание с наружной мембраной органеллы. Как показано на упрощенной диаграмме РИС. 10.34, лидерная последовательность инициирует взаимодействие между предшественником белка и мембраной органеллы. Белок проходит сквозь мембрану, и лидерная последовательность отщепляется протеазой со стороны органеллы.

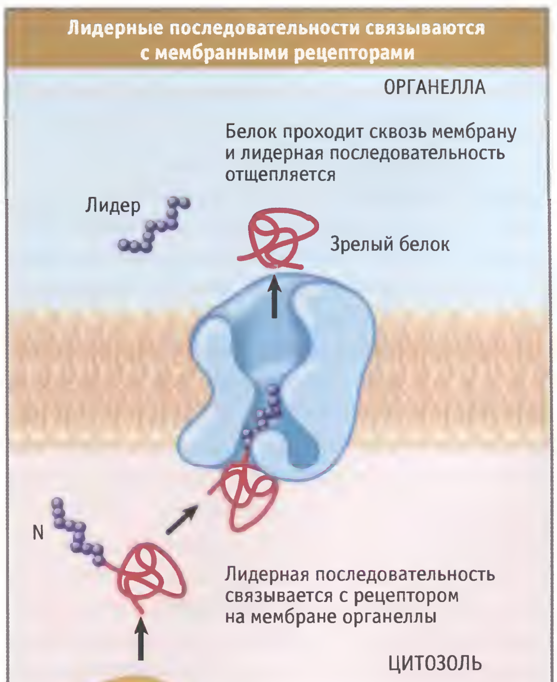


РИС. 10.34 Лидерные последовательности позволяют белкам посттрансляционно распознавать поверхности митохондрий и хлоропластов



РИС. 10.35 Лидерная последовательность субъединицы IV цитохром *c*-оксидазы дрожжей состоит из двадцати пяти нейтральных и основных аминокислот. Первых двенадцати аминокислот достаточно для транспорта любого связанного с ними белка в митохондриальный матрикс

Лидерные последовательности белков, импортируемых в митохондрии и хлоропласты, обычно представлены чередованием гидрофобных и основных аминокислот. Для этих последовательностей характерно практически полное отсутствие кислых аминокислот. Каких-либо общих консервативных последовательностей, свойственных лидерным пептидам органелл, не выявлено. Пример структуры лидера белков органелл приведен на РИС. 10.35. Установлено, что узнавание лидерного пептида зависит не столько от его последовательности, сколько от способности формировать амфипатическую спираль, в которой одна поверхность представлена гидрофобными аминокислотами, а другая — основными.

Лидерная последовательность содержит всю информацию, необходимую для локализации белка органеллы. Это свойство лидерного пептида можно доказать путем конструирования искусственных белков, в которых лидер от белка органеллы соединен с цитоплазматическим белком. В ходе эксперимента создают химерный ген, который в клетке транслируется в химерный белок.

В результате таких экспериментов было обнаружено несколько разных лидерных последовательностей, способных направлять любой присоединенный к ним пептид в митохондрию или хлоропласт. Например, если лидерную последовательность, изображенную на рис. 10.35, присоединить к цитоплазматическому белку ДГФР (дигидрофолатредуктаза), он оказывается локализован в митохондриях.

Лидерная последовательность и транспортируемый белок представляют собой отдельные домены, сворачивающиеся независимо. Следовательно, для любого присоединенного к лидеру белка лидерный пептид должен быть способен принять конформацию, необходимую для узнавания рецепторами мембраны органеллы. Присоединенный белок не участвует в этом распознавании.

Существуют ли какие-либо ограничения на транспорт гидрофильных белков через гидрофобную мембрану? Ответ на этот вопрос был получен в экспериментах с метотрексатом — лигандом фермента ДГФР, блокирующим транспорт в митохондрию меченного лидерным пептидом ДГФР. Оказалось, что связывание с метатрексатом мешает разворачиванию белка в момент транслокации его через мембрану. Последовательность транспортируемого белка не влияет на его ассоциацию с мембраной органеллы, однако для того, чтобы пройти сквозь мембрану вслед за лидерной последовательностью, белок должен обладать пластичностью и способностью к полной денатурации.

Для осуществления процесса транслокации через мембрану, как на внешней, так и на внутренней ее стороне необходим гидролиз АТФ. Этот донор энергии может быть вовлечен и в проталкивание белка снаружи, и в протягивание его с внутренней стороны мембраны. Для импорта белков в митохондрии и экспорта белков из бактериальной клетки необходимым условием переноса терминальной части лидерной последовательности сквозь мембрану (для митохондрий — через внутреннюю мембрану) является наличие электрохимического потенциала.

10.17 Серия лидерных участков определяет детальное местоположение белка внутри органеллы

Ключевые понятия:

- N-терминальная часть лидерной последовательности определяет попадание белка в матрикс митохондрий или внутреннее пространство хлоропластов.
- Последующие участки лидерного пептида определяют дальнейшее встраивание белка в мембрану либо его попадание в межмембранное пространство.
- Лидерные последовательности отщепляются от белка одна за другой.

Митохондрии окружены оболочкой, состоящей из двух мембран. Белки, транспортируемые в митохондрию, могут быть предназначены для работы в наружной мембране, межмембранном пространстве, во внутренней мембране, или в матриксе. Кроме того, трансмембранные белки могут принимать различную ориентацию (выступая с одной или с другой стороны мембраны).

Что определяет попадание митохондриального белка в нужный компартмент органеллы? «По умолчанию»

импортируемый в митохондрию белок должен пройти сквозь обе мембраны в матрикс. Этот путь определяет N-концевая часть лидерной последовательности. Белок, предназначенный для межмембранного пространства или для внутренней мембраны, требует дополнительного сигнала, определяющего его местоположение внутри органеллы. Состоящая из нескольких частей лидерная последовательность содержит сигналы, функционирующие по иерархическому принципу (как показано на **РИС. 10.36**). Первая часть лидерного пептида направляет белок к органелле, вторая часть присутствует в том случае, если конечным пунктом транспортировки белка является не матрикс. Обе части лидерной последовательности удаляются в результате последующего отщепления.

В качестве примера можно привести цитохром *c1*. Связанный с внутренней мембраной митохондрии, этот белок обращен в межмембранное пространство. Его лидерная последовательность состоит из шестидесяти одной аминокислоты и может быть поделена на отдельные регионы, выполняющие разные функции. Первые тридцать две аминокислоты (или даже только N-концевая часть этой последовательности) самостоятельно обеспечивают транспорт ДГФР в матрикс митохондрии. Из этого следует, что первая часть лидерной последовательности (тридцать две N-концевые аминокислоты) обеспечивает сигнал, направляющий белок в матрикс. Однако полная лидерная последовательность транспортирует присоединенный белок — такой, как мышиный ДГФР, — в межмембранное пространство.

Что же не дает белку двигаться сквозь межмембранное пространство, когда он связан с полной лидерной последовательностью? Регион, следующий за матрикс-направляющим сигналом (представляющий собой участок из девятнадцати аминокислот), формирует другой сигнал, отвечающий за локализацию белка во внутренней мембране или в межмембранном пространстве. Для упрощения назовем этот сигнал мембран-направляющим сигналом.

Две части лидерной последовательности, представляющие собой два типа сигналов, имеют разный аминокислотный состав. Как показано на **РИС. 10.37**, тридцать пять N-концевых аминокислот своим высоким содержанием незаряженных аминокислот, перемежающихся основными аминокислотами, напоминают прочие лидерные последовательности, направляющие белки в органеллы. Однако следующие девятнадцать аминокислот представляют собой сплошную последовательность незаряженных аминокислот, достаточно длинную для того, чтобы связаться с липидным бислоем. Мембран-направляющий сигнал напоминает последовательности, участвующие в транслокации белка через мембраны ЭПР (см. разд. 10.7 «Сигнальные последовательности инициируют транслокацию»).

Отщепление матрикс-направляющего сигнала — единственная модификация, которой подвержены белки матрикса. Однако эта же сигнальная последовательность отщепляется и в случае белков, предназначенных для межмембранного пространства, причем после ее отщепления новой N-концевой последовательностью белка становится мембран-направляющий сигнал, который и

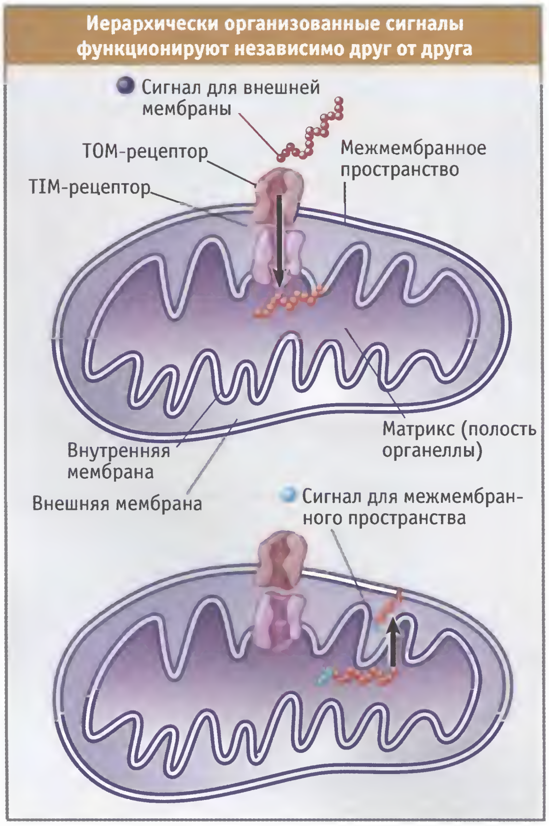


РИС. 10.36 На внешней и внутренней мембранах митохондрий расположены рецепторы для транспорта белков. Распознавание лидерной последовательностью рецепторов внешней мембраны приводит к переносу белка через оба рецепторных комплекса в матрикс, где лидерная последовательность отщепляется. При наличии сигнала, определяющего локализацию в мембране органеллы, белок экспортируется туда уже из матрикса

доставляет белок к его конечному пункту назначения во внешней мембране, межмембранном пространстве или во внутренней мембране органеллы, и только после этого отщепляется.

Похожим образом N-концевой матрикс-направляющий сигнал функционирует и в отношении всех остальных митохондриальных белков. Распознавание его рецепторами внешней мембраны обеспечивает транспорт белка сквозь обе мембраны. Следует отметить, что независимо от конечного пункта назначения белка матрикс-направляющий сигнал отщепляет одна и та же протеаза. Эта протеаза является водорастворимым Mg^{2+}-зависимым ферментом матрикса митохондрий. Следовательно, N-концевая матрикс-направляющая последовательность должна достигнуть матрикса даже в том случае, если сам белок предназначен для межмембранного пространства.

Как уже упоминалось, локализация белка в матриксе определяется отсутствием любого другого сигнала. Мембран-направляющий сигнал активируется в результате отщепления мактрикс-направляющего сигнала в матриксе митохондрий. Оставшаяся часть лидерной после-



РИС. 10.37 Лидерная последовательность дрожжевого цитохрома *c*1 представлена N-концевым участком, направляющим белок к митохондрии, за которым следует участок (отщепляющийся впоследствии), направляющий белок к внутренней мембране. Лидерная последовательность удаляется полностью в результате двух реакций расщепления

довательности (которая теперь является N-концевой) и определяет конечное местоположение белка.

Природа мембран-направляющего сигнала спорна. Согласно одной из моделей, в матрикс проникает целый белок, после чего мембран-направляющий сигнал стимулирует обратный транспорт белка на внутреннюю мембрану или даже сквозь нее. Альтернативная модель предполагает, что мембран-направляющий сигнал не дает большей части белка пройти за лидерной последовательностью сквозь внутреннюю мембрану в матрикс. Какая бы из моделей ни соответствовала действительности, еще одна протеаза (расположенная в межмембранном пространстве) завершает удаление дополнительных лидерных последовательностей.

Аналогичным образом осуществляется транспорт белков сквозь мембраны хлоропласта. РИС. 10.38 демонстрирует разнообразие местоположений белков хлоропласта. Сначала они проходят сквозь внешнюю и внутреннюю мембраны хлоропласта в строму (по тому же принципу, что и белки в матрикс митохондрий). Некоторые белки транспортируются дальше на мембрану тилакоидов или в просвет тилакоидов (очевидно, что направляющиеся к тилакоидным мембранам или в просвет тилакоидов белки должны прежде пересечь строму).

Сигналы, направляющие белок в хлоропласт, напоминают сигналы, направляющие белок в митохондрию.



РИС. 10.38 Белок транспортируется в хлоропласт с помощью лидерной последовательности длиной ~50 аминокислот. N-концевая часть лидерной последовательности определяет попадание белка в оболочку хлоропласта либо сквозь нее в строму. Отщепление лидерной последовательности происходит во время транспорта через оболочку хлоропласта

Лидерная последовательность состоит из ~50 аминокислот, и ее N-концевая часть необходима для распознавания оболочки хлоропласта. После попадания белка в строму происходит расщепление лидерной последовательности между 20-й и 25-й позициями, в результате обнажая новый N-концевой лидер, обеспечивающий узнавание тилакоидной мембраны или просвета тилакоидов. В хлоропластах существуют по меньшей мере четыре разные системы, катализирующие импорт белков в тилакоидную мембрану.

Общий принцип, управляющий транспортом белка в митохондрии и хлоропласты, заключается в том, что N-концевая часть лидерной последовательности направляет белок в матрикс органеллы, а дополнительная (вторая) часть лидера необходима для локализации белка в наружной мембране, межмембранном пространстве или на внутренней мембране органеллы.

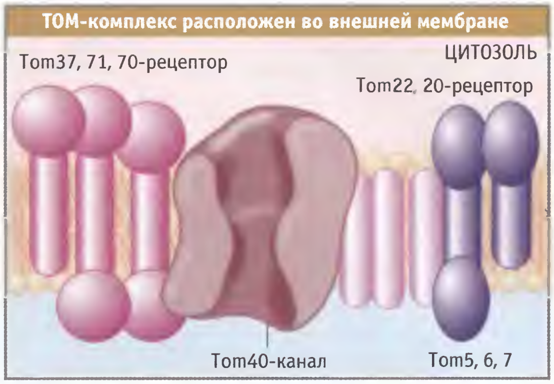


РИС. 10.39 ТОМ-белки формируют рецепторный комплекс (комплексы), необходимый для транслокации через внешнюю мембрану митохондрий

10.18 Внешняя и внутренняя мембраны митохондрий содержат разные транслоконы

Ключевые понятия:

- Транспорт белков через внешнюю и внутреннюю мембраны осуществляется при участии разных рецепторных комплексов.
- Расположенный на внешней мембране комплекс ТОМ представляет собой крупную белковую структуру, направляющую транспортируемые белки в канал Tom40 при помощи одного из двух субкомплексов.
- В зависимости от того, куда должен попасть транспортируемый белок (на внутреннюю мембрану или в матрикс), на внутренней мембране митохондрий для транспорта используются разные TIM-комплексы.
- Белки переходят от комплекса ТОМ непосредственно к комплексу TIM.

Транспорт через каждую из мембран оболочки хлоропласта или митохондрии опосредован разными рецепторами. Рецепторы внешней и внутренней мембран хлоропласта называют TOC и TIC, а рецепторы митохондрии — **TOM** и **TIM** соответственно.

Комплекс ТОМ состоит из ~9 белков, многие из которых являются трансмембранными. Общая модель комплекса показана на РИС. 10.39. Он имеет размер >500 кДа, диаметр ~138 Å и формирует ион-проводящий канал. Комплекс содержит 2–3 отдельных кольца диаметром 75 Å, в каждом из которых есть пора диаметром 20 Å.

Белок Tom40 пронизывает мембрану, формируя собственно канал для транслокации. Он контактирует с пре-белками в тот момент, когда они проходят через внешнюю мембрану митохондрии. Tom40 связан с тремя сравнительно небольшими белками — Tom5, 6, и 7, которые могут быть компонентами канала или факторами сборки. Два дополнительных субкомплекса представляют собой поверхностные рецепторы. Белки Tom20 и Tom22 вместе образуют субкомплекс, экспонирующий свои домены в цитоплазму. Большинство белков, импортируемых в митохондрию, узнаются именно субкомплексом Tom20,22, являющимся первичным рецептором, распознающим N-концевую последовательность транслоцируемого белка. Белки Tom37,70,71 образуют рецептор, связывающий лишь небольшое число белков, имеющих внутреннюю сигнальную последовательность.

Проходящий сквозь ТОМ-комплекс белок, как правило, не высвобождается в межмембранное пространство, а переносится непосредственно в TIM-комплекс. В экспериментах можно даже обнаружить промежуточную форму транслоцируемого белка, в которой лидерная последовательность уже удалена протеазами матрикса, а большая часть предшественника белка все еще экспонирована с цитоплазматической стороны внешней мембраны. Основываясь на этих данных, можно утверждать, что в момент транслокации белок пронизывает сразу обе митохондриальные мембраны. Комплексы ТОМ и TIM не взаимодействуют друг с другом напрямую (по крайней мере, они не формируют детектируемого стабильного комплекса). По всей видимости, они просто связываются белком в процессе его перемещения. Как только транслоцируемый белок достигает межмембран-



РИС. 10.40 Белки Tim формируют рецепторный комплекс для транслокации через внутреннюю мембрану митохондрии

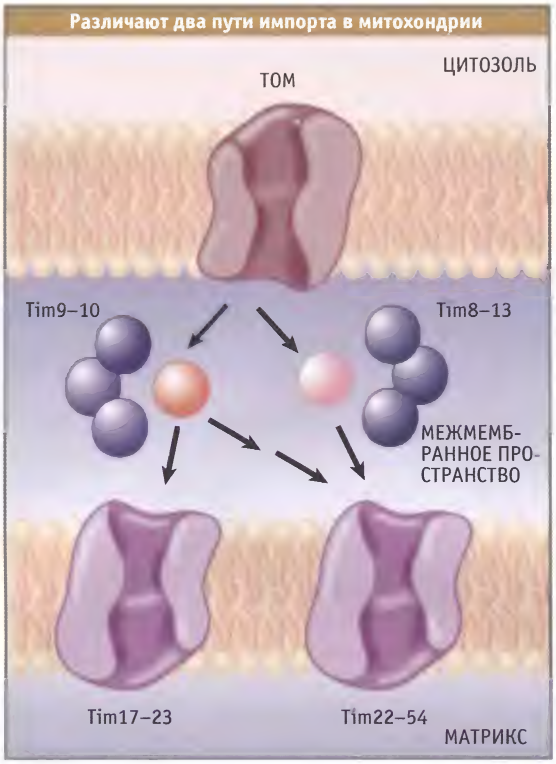


РИС. 10.41 Tim9–10 направляют белки от ТОМ-комплекса к любому из TIM-комплексов, Tim8–13 — только к Tim22–54

ного пространства, экспонируемые им аминокислотные остатки могут немедленно связываться с комплексом TIM, в то время как оставшаяся часть белка продолжает транслоцироваться через ТОМ-комплекс.

Во внутренней мембране митохондрий находятся два разных TIM-комплекса. Комплекс Tim17–23 переносит белки в просвет органеллы. Причем узнавание субстратного полипептида идет по наличию у него положительно заряженного N-концевого сигнала. Трансмембранные белки Tim17-Tim23 формируют канал. **РИС.** 10.40 демонстрирует, как они связаны с белком Tim44 на матриксной стороне мембраны. Tim44, в свою очередь, связывает шаперон Hsp70, образующий комплекс с шапероном Mge (аналогом бактериального GrpE). Эти взаимодействия гарантируют, что в момент вхождения импортируемого белка в матрикс он свяжется шапероном Hsp70. Высокая аффинность связывания Hsp70 с белком, выходящим из внутренней мембраны в развернутой конформации, помогает «протянуть» белок сквозь канал.

Основная шаперонная активность в митохондриальном матриксе принадлежит Hsp60 (формирующему такую же структуру, что и его аналог GroEL). Ассоциация белка с Hsp60 необходима для соединения субъединиц импортируемого белка, формирующих олигомерный комплекс. Белок может переходить от Hsp70 к Hsp60 в процессе приобретения им зрелой конформации.

Второй TIM-комплекс, Tim22-54, переносит белки, предназначенные для внутренней мембраны митохондрий.

Как транспортируемый белок находит путь от ТОМ-комплекса к необходимому TIM-комплексу? Два белковых комплекса в межмембранном пространстве сопровождают транслоцируемый белок от TIM к ТОМ. Комплексы Tim9-10 и Tim8-13 служат как бы проводниками для разных групп субстратных белков. Tim9-10 направляет белки к Tim22-54 или Tim23-17, в то время как Tim8-13 — только к Tim22-54. Некоторые транслоцируемые белки не используют ни Tim9-10, ни Tim8-13, следовательно, должны существовать какие-то другие пути транспорта. Перечисленные варианты проиллюстрированы на **РИС.** 10.41.

Какую роль играют сопровождающие комплексы? Они могут способствовать выходу белка из ТОМ-комплекса, а также узнаванию его TIM-комплексом. Как показано на **РИС.** 10.42, транслоцируемый белок может пройти напрямую из ТОМ-канала к Tim9-10-комплексу, а затем — в канал Tim22-54.

Формирование пространственной структуры митохондриальных белков идет в разных условиях в момент синтеза на рибосоме и после прохождения через мембраны. Цитоплазма и митохондриальный матрикс различаются по ионному окружению и имеющимся шаперонам. Возможно, митохондриальный белок способен достичь своей зрелой конформации *только* внутри митохондрии.

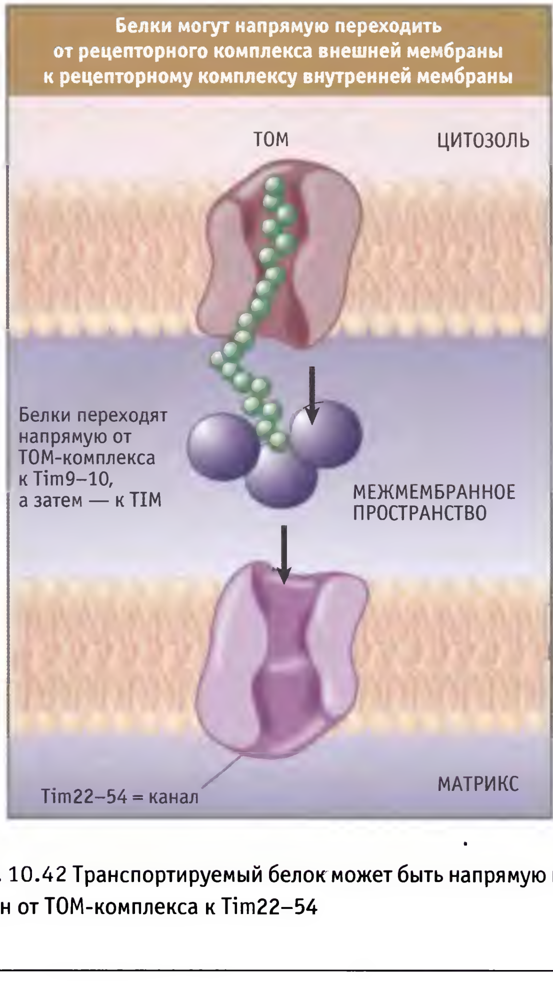


РИС. 10.42 Транспортируемый белок может быть напрямую перенесен от ТОМ-комплекса к Tim22–54

10.19 Пероксисомы обладают собственной системой транслокации белков

Ключевые понятия:

- Белки импортируются в пероксисомы в виде молекул с правильной пространственной структурой.
- Они имеют либо последовательность PTS1 на C-конце молекулы, либо последовательность PTS2 на N-конце.
- Рецептор Pex5p связывает последовательность PTS1, рецептор Pex7p связывает последовательность PTS2.
- Рецепторы представляют собой цитоплазматические белки, перемещающиеся в пероксисому в комплексе с транспортируемым белком и затем возвращающиеся в обратно в цитоплазму.

Пероксисомы представляют собой небольшие внутриклеточные тельца (от 0,5 до 1,5 нм в диаметре), ограниченные единственной мембраной. Они содержат ферменты, ответственные за утилизацию кислорода и превращающие кислород в перекись водорода путем отщепления атомов водорода от различных субстратов. Фермент каталаза затем использует перекись водорода для окисления множества других субстратов. Активность этих ферментов имеет для клетки критическое значение. С тех пор, как было обнаружено, что смертельное заболевание — синдром Зельвегера — вызывается дефицитом пероксисом, еще для более чем 15 заболеваний человека была показана ассоциация с расстройством функции пероксисом.

Все компоненты пероксисом импортируются из цитоплазмы. Белки, необходимые для формирования пероксисом, называются **пероксинами.** Обнаружено двадцать три гена, кодирующих пероксины, а пероксисомные болезни человека разделены на двенадцать групп, и для большинства из них показана ассоциация с конкретным геном пероксина. У клеток, несущих мутации, полностью выключающие некоторые из этих генов, пероксисомы отсутствуют. В отдельных случаях введение в такую клетку генов дикого типа приводит к возникновению пероксисом. Существует гипотеза, согласно которой пероксисомы возникают только в результате дупликации уже существующих пероксисом. Для ее подтверждения или опровержения необходимо ответить на вопрос: могут ли пероксисомы быть собраны *de novo* из отдельных компонентов? По крайней мере в некоторых случаях даже при отсутствии пероксинов в клетке сохраняются пероксисомные «тени» — пустые мембранные тельца. И даже в том случае, если в клетке не удается обнаружить каких-либо остаточных телец (служащих для регенерации пероксисом), доказать их полное отсутствие крайне сложно.

Транспорт белков в пероксисомы осуществляется посттрансляционно. Импортируемые внутрь пероксисом полипептиды несут одну из двух коротких последовательностей, называемых пероксисомо-направляющим сигналом (PTS1 и PTS2). Сигнал PTS1 представляет собой три- или тетрапептид, расположенный на C-конце молекулы. Первоначально он был охарактеризован как последовательность SKL (серин-лизин-лейцин), но к настоящему времени показано, что различные три- и тетрапептиды могут функционировать в качестве сигнала PTS1. Добавления соответствующей последовательности к C-концу цитоплазматического белка достаточно для того, чтобы он был импортирован в пероксисому. Сигнал PTS2 представляет собой последовательность из девяти аминокислот (также характеризующуюся множеством вариантов). Эта последовательность может располагаться вблизи N-конца белка или же быть внутренней. Возможно, существует и третий тип меток — PTS3.

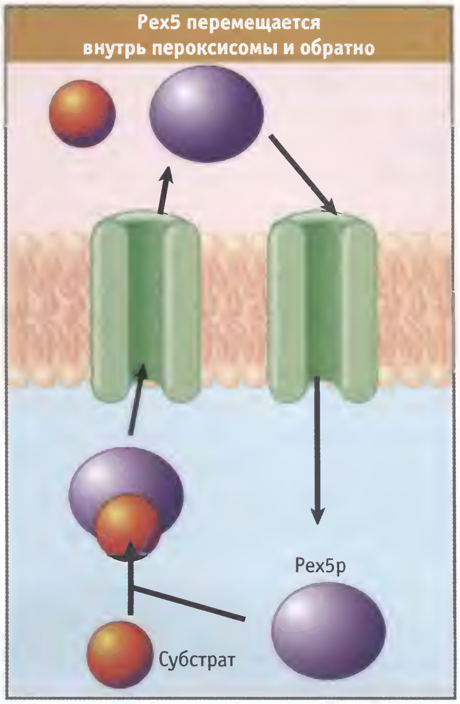


РИС. 10.43 Рецептор Pex5 связывает субстратный белок в цитозоле, переносит его через мембрану в пероксисому, а затем возвращается в цитозоль

Несколько пероксисомных белков необходимы для импорта белков из цитоплазмы. Пероксисомные рецепторы, связывающие два обозначенных выше типа сигналов, называются Pex5p и Pex7p соответственно. Другие белки являются частью трансмембранного комплекса, участвующего в реакции транслокации.

Транспорт в пероксисомы характеризуется несколькими необычными чертами, подчеркивающими его отличие от систем транспорта белков в другие органеллы.

Белки могут быть импортированы в пероксисомы в своем зрелом, полностью свернутом состоянии, что невозможно при транспорте белков в ЭПР или митохондрии (сквозь мембранные каналы которых белок проходит в виде развернутой нити аминокислот). Не вполне понятно, каким образом структура трансмембранного канала может растянуться настолько, чтобы пропустить зрелый белок. Одним из возможных объяснений является идея о сборке канала непосредственно вокруг транспортируемого белка в тот момент, когда он ассоциирует с мембраной.

Рецепторы Pex5p и Pex7p не являются мембрансвязанными белками, большая их часть расположена в цитоплазме (и лишь небольшая часть ассоциирована с пероксисомами). Рецепторы циркулируют между пероксисомой и цитоплазмой. На РИС. 10.43 показано, как рецептор связывает субстратный белок в цитоплазме, переносит его к пероксисоме, проходит с ним через мембрану во внутреннюю среду, а затем возвращается обратно в цитоплазму для повторения цикла. Такая «челночная» схема работы рецептора напоминает систему переносчиков, осуществляющих импорт белков в ядро.

На мембране пероксисомы и Pex5p, и Pex7p взаимодействуют с одним и тем же мембранным белковым комплексом, состоящим из Pex14p и Pex13p. Рецепторы связываются с этим комплексом, после чего несколько дополнительных пероксинов вовлекаются в процесс трансмембранного переноса. Детали этого процесса до сих пор не вполне ясны.

Белки, встраивающиеся в мембрану пероксисомы, имеют последовательность, называемую mPTS, однако о процессе их интеграции известно мало. Pex3p, возможно, является ключевым белком встраивания в мембрану, поскольку в его отсутствие другие белки в мембранах пероксисом не обнаруживаются. Pex3p имеет свою собственную последовательность mPTS, что поднимает вопрос о механизме встраивания в мембрану самого Pex3p. Возможно, mPTS новых Pex3p взаимодействует с уже находящимися в мембране Pex3p. Это возвращает нас к вопросу о том, могут ли пероксисомы быть собраны *de novo.*

10.20 Бактерии используют как котрансляционную, так и посттрансляционную системы транспорта белков

Ключевые понятия:

- Для транспорта бактериальных белков на мембрану или сквозь нее используются оба механизма транслокации: котрансляционный и посттрансляционный.

Бактериальная оболочка (грамотрицательных бактерий) состоит из двух мембран, пространство между которыми называют **периплазмой.** Белки, экспортируемые из цитоплазмы, могут оставаться в составе бактериальной оболочки или секретироваться клеткой вовне. Механизм бактериальной секреции схож с механизмами, описанными для эукариотических клеток: в обеих системах можно отметить гомологичные компоненты. Как изображено на РИС. 10.44, дальнейшая судьба белков, экспортируемых из цитоплазмы, может развиваться по одному из четырех путей:

- встраивание во внутреннюю мембрану;
- перенос через внутреннюю мембрану в периплазму;
- встраивание во внешнюю мембрану;
- перенос через внешнюю мембрану во внешнюю среду.

На внутренней мембране бактериальной оболочки разные белковые комплексы отвечают за транспорт белков, проходящих сквозь мембрану, и белков, интегрирующих в мембрану. Это напоминает ситуацию с митохондриями, где разные комплексы (как во внутренней, так и во внешней мембране) взаимодействуют с разными группами субстратных белков в зависимости от их конечного пункта назначения (см. разд. 10.16 «Посттрансляционная интеграция белков в мембрану зависит от лидерных последовательностей»). Разница между экспортом из бактериальной клетки и импортом в органеллы заключается в том, что бактериальный белковый транспорт может быть как ко-, так и посттрансляционным. Некоторые белки секретируются и ко-, и посттрансляционно (при этом баланс двух процессов определяется относительной скоростью трансляции в сравнении со скоростью переноса через мембрану).

Экспортируемые бактериальные белки имеют N-концевую лидерную последовательность с гидрофильным N-концом и прилегающим к нему гидрофобным ядром. Лидерная последовательность отщепляется сигнальной пептидазой, узнающей предшественников нескольких экспортируемых белков. Сигнальная пептидаза является мембранным белком, расположенным на внутренней мембране оболочки бактерий. Мутации в N-концевой лидерной последовательности нарушают секрецию, однако они супрессируются мутациями в генах, кодирующих компоненты аппарата белкового транспорта. Несколько генов под общим названием *sec* кодируют компоненты секреторного аппарата. Мутации в них блокируют секрецию многих (или даже всех) экспортируемых белков.



РИС. 10.44 Бактериальные белки могут транспортироваться через плазматическую мембрану пост- и котрансляционно. Конечным местом их локализации могут быть внутренняя или внешняя мембраны, межмембранное пространство, либо они могут быть секретированы

10.21 Sec-система транспортирует белки на внутреннюю мембрану бактериальной оболочки и сквозь нее

Ключевые понятия:

- Бактериальный транслокон SecYEG соответствует транслокону Sec61 у эукариот.
- В процесс попадания транспортируемого белка в транслокон вовлечены различные шапероны.

Существует несколько систем транспорта белков через внутреннюю мембрану оболочки бактерий. Лучше всего охарактеризована система Sec, компоненты которой изображены на РИС. 10.45. Погруженный в мембрану транслокон состоит из трех субъединиц, аналогичных компонентам комплекса Sec61 у млекопитающих и дрожжей. Каждая из субъединиц является трансмембранным белком (SecY обладает десятью трансмембранными сегментами, SecE — тремя). Функционирующий транслокон представляет собой тример, содержащий по одной субъединице каждого типа. Главный путь, по которому белки направляются в транслокон, опосредован белками SecA и SecB. SecB представляет собой шаперон, связывающий белок в процессе синтеза и контролирующий его фолдинг. SecB переносит белок к SecA, который, в свою очередь, транспортирует белок на транслокон.

Как показано на РИС. 10.46, существует два основных пути доставки белка в канал Sec:

- шаперон SecB, и
- SRP на основе 4,5S RNA.

Несколько шаперонов могут повышать эффективность экспорта бактериальных белков, предотвращая их преждевременное сворачивание; среди них «фактор запуска» (trigger factor) (охарактеризованный как шаперон, принимающий участие в экспорте белков), GroEL (см. разд. 10.5 «Шапероны необходимы для формирования структуры только что синтезированных или денатурированных белков» и разд. 10.18 «Внутренняя и внешняя митохондриальные мембраны обладают разными транслоконами»), а также SecB (идентифицированный как продукт одного из мутантов *sec*). Из перечисленных факторов SecB наименее представлен, однако он играет главную роль в обеспечении экспорта белков. Его роль заключается в выполнении двух функций. Во-первых, SecB действует как шаперон, связывая растущий полипептид и тормозя его фолдинг. Он не может денатурировать структуру уже свернувшегося белка, поэтому не способен выступать в качестве денатурирующего фактора. В этой связи его задача состоит в ингибировании неправильного сворачивания синтезируемого белка. Во-вторых, SecB характеризуется сродством к белку SecA, что позволяет ему направлять предшественник белка к внутренней мембране бактериальной клетки. Представленный на РИС. 10.47 путь SecB-SecYEG используется для транслокации белков, секретируемых в периплазму.



РИС. 10.45 Система Sec представляет собой SecYEG транслокон, встроенный в мембрану: компонент SecA проталкивает транспортируемые белки через канал, SecB — шаперон, который переносит растущую полипептидную цепь к SecA; сигнальная пептидаза отщепляет N-концевую сигнальную последовательность от переносимого через мембрану белка

SecA — большой примембранный белок, взаимодействующий с внутренней мембраной несколькими способами. Как и большинство примембранных белков, он связан с мембраной благодаря своей аффин-



РИС. 10.46 Система SecB/SecA переносит белки к транслокону, пронизывающему мембрану. 4,5S РНК осуществляет перенос белков, которые затем остаются в мембране

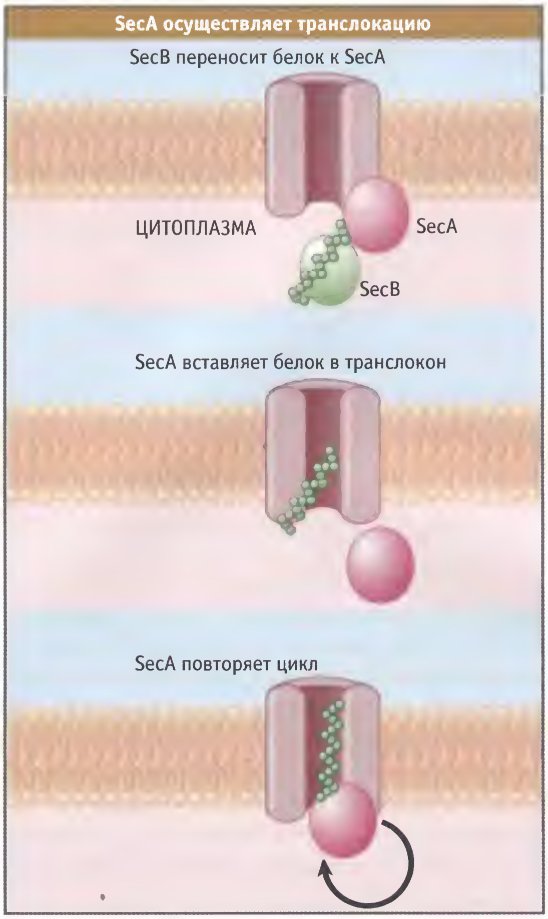


РИС. 10.47 **SecB переносит растущий пептид к SecA, который встраивает белок в канал. Для реализации этого процесса необходима энергия гидролиза АТФ и протон-движущая сила. SecA ассоциирует с каналом и диссоциирует с него, обеспечивая движущую силу для проталкивания белка сквозь канал**

ности к кислым липидам. Кроме того, он обладает сродством к SecY-элементу транслокона, входящего в состав многокомпонентного комплекса, обеспечивающего процесс транслокации. Однако в присутствии дополнительных белков (SecD и SecF) SecA может временно диссоциировать от мембраны. Возможно, SecA является своеобразным «двигателем», с помощью которого субстратный белок проталкивается через SecYEG транслокон.

SecA взаимодействует и с SecB, и с белком-предшественником (для которого SecB является шапероном). Скорее всего, для этого взаимодействия необходимо, чтобы белок имел правильную пространственную структуру и лидерный пептид. SecA обладает АТФазной активностью, зависящей от связывания его с липидами, SecY и предшественником белка. Во время транслокации АТФаза функционирует в циклическом режиме. После того как SecA связался с предшественником белка, он связывает АТФ и проталкивает ~20 аминокислот сквозь мембрану. Для диссоциации SecA от транспортируемого белка необходим гидролиз АТФ. Затем этот цикл повторяется: SecA снова ассоциирует с белком, связывает новую молекулу АТФ, транслоцирует очередной сегмент белка и диссоциирует. В ходе транслокации SecA может переходить из примембранной во внутримембранную форму (во время каждого цикла 30 кДа домен SecA погружается в мембрану и извлекается из нее).

Электрохимический потенциал на мембране может содействовать транслокации белков. Когда SecA диссоциирует от белка, тот может «протягиваться» сквозь мембрану протон-движущей силой. Сам по себе электрохимический потенциал не может инициировать перенос белка через мембрану, но он может способствовать процессу, запущенному в результате циклической работы АТФазы SecA. Таким образом, после (или между) циклами действия SecA-АТФ протондвижущая сила может продолжать транслокацию предшественника белка.

Рибонуклеопротеиновый комплекс 4,5S РНК с белками Ffh и FtsY у *E. coli* является аналогом эукариотической SRP (см. разд. 10.9 «SRP взаимодействует с SRP-рецептором»). Возможно, этот комплекс играет роль в поддержании необходимой конформации синтезируемого белка до тех пор, пока он не провзаимодействует с другими компонентами секреторного аппарата. Участие Ffh и FtsY показано для секреции некоторых (но не всех) белков у *E.coli*. Как можно видеть на рис. 10.46, субстратом «бактериальной SRP» являются трансмембранные белки. Основой для избирательного связывания лишь некоторых экспортируемых белков является распознавание SRP *E.coli* якорной последовательности (по определению, якорная последовательность присутствует только в мембранных белках). Хлоропласты обладают белками, аналогичными Ffh и FtsY, которые, однако, не связаны с РНК-компонентом.

10.22 Sec-независимая система транспорта белков *E. coli*

Ключевые понятия:

- Системы для транслокации белков через клеточную мембрану *E. coli* сходны с системами, характерными для органелл.
- Одна из систем позволяет определенным белкам встраиваться в мембрану без участия аппарата транслокации
- YidC гомологичен митохондриальной системе переноса белков во внутреннюю мембрану.
- Tat-система транспортирует белки с двойным аргининовым мотивом в периплазматическое пространство.

Удивительная альтернативная система транслокации белков была открыта у *E.coli* при изучении поверхностного белка фага M13. Как показано на РИС. 10.48,

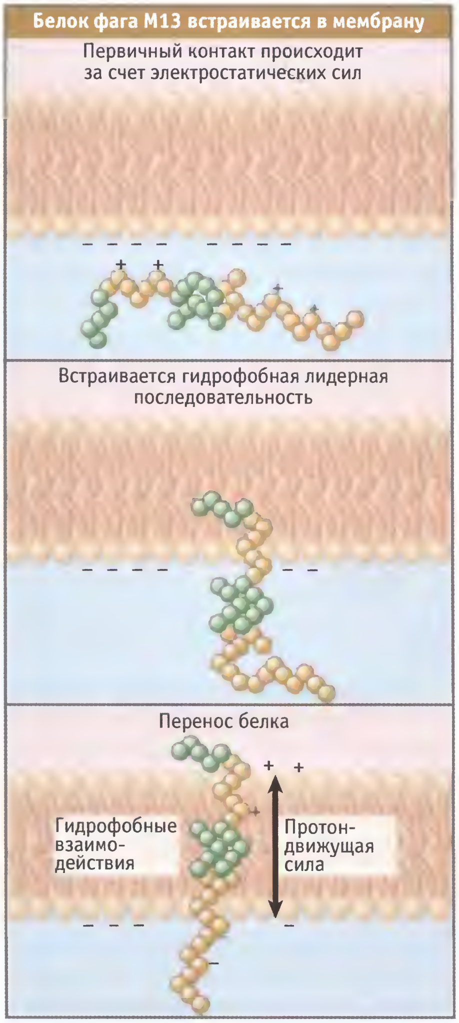


РИС. 10.48 Поверхностный белок фага M13 сначала образует контакты с мембраной за счет электростатических сил, а затем встраивает в нее свои гидрофобные последовательности. Транслокация за счет гидрофобных взаимодействий и протон-движущей силы продолжается до тех пор, пока якорная последовательность не войдет в мембрану

перенос этого белка через мембрану вовсе не требует наличия аппарата транслокации! Посттрансляционно он может проникать в безбелковые липосомы. Начальное взаимодействие белка с мембраной определяется специальными (состоящими из основных аминокислот) последовательностями в N- и C-концевых регионах белка. Эти регионы могут взаимодействовать с отрицательно заряженными головками фосфолипидов. Затем белок проникает в мембрану, используя гидрофобные группы в N-концевой лидерной последовательности и аминокислоты внутренней якорной последовательности. Гидрофобность является основной движущей силой транслокации, однако дополнительный вклад может вносить и протон-движущая сила, создаваемая между положительно заряженной периплазматической стороной мембраны и кислыми регионами белка. Описанный механизм ведет к самостоятельному проникновению белка сквозь мембрану, после чего пептидаза отщепляет N-концевую лидерную последовательность. Наличие такого механизма для других бактериальных белков не установлено; возможно, он применим только к поверхностному белку бактериофага, однако некоторые белки хлоропласта похожим способом могут встраиваться в мембрану тилакоида.

Мутации в гене *yidC* блокируют встраивание белков во внутреннюю мембрану бактерий. YidC является гомологом белка Oxa1p, необходимого для встраивания белков из матрикса во внутреннюю мембрану митохондрий. Он может функционировать как независимо от SecYEG, так и совместно с ним. Интеграция некоторых YidC-зависимых белков требует участия SecYEG. Это означает, что YidC (совместно с транслоконом) направляет процесс в сторону встраивания белка в мембрану, а не его секреции. Однако некоторые белки, процесс интеграции которых в мембрану также зависит от YidC, не требуют участия SecYEG: возможно, что в этом случае вместо выполнения функций транслокона требуется выполнение каких-то других (еще не установленных) функций.

Tat (twin arginine targeting)-система (названная так за свою способность транспортировать белки, несущие два аргинина подряд) ответственна за транслокацию белков, тесно связанных с кофакторами (поскольку такие белки имеют ограниченную способность к денатурации во время транспорта через мембрану). Этот подход противоречит механизму большинства систем транслокации, в которых белки переносятся через мембрану в развернутом состоянии и только после прохождения мембраны сворачиваются в зрелую конформацию (схожим образом функционирует система Hcf106, осуществляющая транспорт белков в просвет тилакоида хлоропласта). Обе описанных выше системы транспортируют белки в периплазму.

10.23 Резюме

Встраивающийся или транспортируемый сквозь мембрану белок имеет сигнальную последовательность, распознаваемую трансмембранными или примембранными рецепторами. Белок транспортируется сквозь гидрофобный канал, формируемый трансмембранным белком (белками). Почти во всех случаях белок проходит сквозь канал в развернутом состоянии, вследствие чего сразу после транслокации ему необходимо связаться с шаперонами для принятия правильной конформации. Важным исключением являются пероксисомы, импорт белков в которые идет без денатурации при помощи цитоплазматического белка, переносящего зрелые белки сквозь канал в мембране.

Синтез белков в цитоплазме начинается на свободных рибосомах. Белки, секретируемые из клетки или встраиваемые в мембрану ретикулоэндотелиальной системы, имеют N-концевую сигнальную последовательность, обеспечивающую присоединение рибосомы к мембране эндоплазматического ретикулума. Перенос белка через мембрану осуществляется котрансляционно. Процесс начинается со связывания сигнальной последовательности сигнал-узнающей частицей SRP (рибонуклеопротеиновая частица), которая останавливает трансляцию. Далее SRP связывается с SRP-рецептором на мембране ЭПР и переносит сигнальную последовательность к мембранному каналу Sec61/TRAM. Синтез возобновляется, и белок проталкивается сквозь мембрану по мере синтеза. Мембранный канал (большей частью сформированный белком Sec61) обеспечивает гидрофильную среду для транслокации.

Секреторный белок целиком проходит сквозь мембрану в просвет ЭПР. Белки, интегрированные в мембраны, согласно их ориентации в мембране делят на две основных группы. У мембранных белков I группы N-концевая сигнальная последовательность отщепляется, а перенос через мембрану останавливает специальная якорная последовательность. Причем белки оказываются ориентированы в мембране таким образом, что их N-концы обращены в ЭПР, а C-концы — в цитоплазму. Белки II группы не имеют отщепляемого N-концевого сигнала. Вместо этого они обладают комбинированной сигнально-якорной последовательностью, которая входит в мембрану и остается погруженной в нее. C-концы таких белков обращены в просвет ЭПР, а N-концы остаются в цитоплазме. Ориентация сигнально-якорной последовательности определяется правилом «положительное внутри», которое гласит, что конец якорной последовательности с наибольшим положительным зарядом будет расположен в цитоплазме. Белки, имеющие единственный трансмембранный регион, могут перемещаться латерально из канала в липидный бислой. Однако белки могут иметь и множественные трансмембранные регионы, образующие петли, выступающие с обеих сторон мембраны. Механизм встраивания в мембрану множественных сегментов пока неизвестен.

В отсутствие сигнала для синтеза в ЭПР синтезированный на рибосоме белок высвобождается в цитоплазму. В митохондрии и хлоропласты белки импортируются уже посттрансляционно. Такие белки имеют N-концевую лидерную последовательность, направляющую их к внешней мембране органеллы, после чего они транспортируются через внешнюю и внутреннюю мембраны в матрикс. Транслокация требует участия АТФ с обязательным прохождением любого белка через внутреннюю мембрану в матрикс органеллы, где N-концевая лидерная последовательность отщепляется протеазой. Мембрансвязанные белки органелл или белки, предназначенные для межмембранного пространства, имеют дополнительный сигнал (который становится N-концевым после отщепления протеазой первой части лидерной последовательности). Этот сигнал обеспечивает экспорт из матрикса к конечному местоположению (либо, по другой гипотезе, останавливает перенос белка до того момента, как он полностью пройдет в матрикс). Контроль сворачивания белка, осуществляемый в митохондриальном матриксе шаперонами Hsp70 и Hsp60, является важной чертой процесса.

Митохондрии и хлоропласты имеют отдельные рецепторные комплексы, формирующие каналы через внешнюю и внутреннюю мембраны. Все импортируемые белки напрямую переходят от TOM-комплекса во внешней мембране к TIM-комплексу во внутренней мембране. Белки, остающиеся в межмембранном пространстве или во внешней мембране, перенаправляются туда от TIM-комплекса уже после того, как проникнут в матрикс. В зависимости от того, имеют ли импортируемые белки N-концевую или внутреннюю сигнальную последовательность, TOM-комплекс использует различные рецепторы, однако оба типа белков он направляет в канал Tom40. На внутренней мембране расположено два типа TIM-рецепторов: один используется для белков, конечным пунктом назначения которых является матрикс; другой — для белков, возвращающихся в межмембранное пространство или на внешнюю мембрану органеллы.

Несмотря на то что бактерии обладают компонентами котрансляционного трансмембранного переноса белков (аналогичными таковым у эукариот), транслокация у бактерий происходит и посттрансляционно. Бактериальный комплекс SecE/Y является транслоказой, а белок SecA, взаимодействуя с каналом, обеспечивает встраивание и продвижение белка сквозь мембрану. Белок SecB является шапероном, переносящим белок к трансмембранному каналу. Некоторые трансмембранные белки входят в канал посредством взаимодействия со структурой, напоминающей эукариотическую SRP и состоящей из 4,5S РНК и белков Ffh и FstY. Белок YidC является гомологом митохондриального белка, необходимого для встраивания некоторых трансмембранных белков в мембрану.

Литература

10.4 Шапероны обеспечивают формирование правильной пространственной структуры белка

Обзор

Ellis, R. J. and van der Vies, S. M. (1991). Molecular chaperones. *Annu. Rev. Biochem.* 60, 321–347.

Fersht, A. R. and Daggett, V. (2002). Protein folding and unfolding at atomic resolution. *Cell* 108, 573–582.

Hartl, F. U. and Hayer-Hartl, M. (2002). Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* 295, 1852–1858.

Исследование

Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science* 181, 223–230.

van den Berg, B., Ellis, R. J., and Dobson, C. M. (1999). Effects of macromolecular crowding on protein folding and aggregation. *EMBO J.* 18, 6927–6933.

10.5 Шапероны необходимы и для вновь синтезированных, и для денатурированных белков

Обзор

Frydman, J. (2001). Folding of newly translated proteins *in vitro*: the role of molecular chaperones. *Annu. Rev. Biochem.* 70, 603–647.

Moarefi, I. and Hartl, F. U. (2001). Hsp90: a specialized but essential protein-folding tool. *J. Cell Biol.* 154, 267–273.

Исследование

Queitsch, C., Sangster, T. A., and Lindquist, S. (2002). Hsp90 as a capacitor of phenotypic variation. *Nature* 417, 618–624.

Rutherford, S. L. and Lindquist, S. (1998). Hsp90 as a capacitor for morphological evolution. *Nature* 396, 336–342.

10.6 Белки семейства Hsp70 встречаются во всех компартментах клетки

Обзор

Bukau, B. and Horwich, A. L. (1998). The Hsp70 and Hsp60 chaperone machines. *Cell* 92, 351–366.

Frydman, J. (2001). Folding of newly translated proteins *in vitro*: the role of molecular chaperones. *Annu. Rev. Biochem.* 70, 603–647.

Georgopoulos, C. and Welch, W. J. (1993). Role of the major heat shock proteins as molecular chaperones. *Annu. Rev. Cell Biol.* 9, 601–634.

Hartl, F. U. (1966). Molecular chaperones in cellular protein folding. *Nature* 381, 571–580.

Исследование

Blond-Elguindi, S., Cwirla, S. E., Dower, W. J., Lipshutz, R. J., Sprang, S. R., Sambrook, J. F., and Gething, M. J. (1993). Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of BiP. *Cell* 75, 717–728.

Flaherty, K. M., DeLuca-Flaherty, C., and McKay, D. B. (1990). Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. *Nature* 346, 623–628.

Flynn, G. C., Pohl, J., Flocco, M. T., and Rothman, J. E. (1991). Peptide-binding specificity of the molecular chaperone BiP. *Nature* 353, 726–730.

Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E., and Hendrickson, W. A. (1996). Structural analysis of substrate binding by the molecular chaperone DnaK. *Science* 272, 1606–1614.

10.7 Сигнальные последовательности инициируют транслокацию

Обзор

Lee, C. and Beckwith, J. (1986). Cotranslational and posttranslational protein translocation in prokaryotic systems. *Annu. Rev. Cell Biol.* 2, 315–336.

Palade, G. (1975). Intracellular aspects of the process of protein synthesis. *Science* 189, 347–358.

Исследование

Blobel, G. and Dobberstein, B. (1975). Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. *J. Cell Biol.* 67, 835–851.

Lingappa, V. R., Chaidez, J., Yost, C. S., and Hedgpeth, J. (1984). Determinants for protein localization: beta-lactamase signal sequence directs globin across microsomal membranes. *Proc. Natl. Acad. Sci. USA* 81, 456–460.

von Heijne, G. (1985). Signal sequences. The limits of variation. *J. Mol. Biol.* 184, 99–105.

10.8 Сигнальные последовательности взаимодействуют с сигнал-узнающими частицами (SRP)

Обзор

Walter, P. and Johnson, A. E. (1994). Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane. *Annu. Rev. Cell Biol.* 10, 87–119.

Исследование

Tjalsma, H., Bolhuis, A., van Roosmalen, M. L., Wiegert, T., Schumann, W., Broekhuizen, C. P., Quax, W. J., Venema, G., Bron, S., and van Dijl, M. (1998). Functional analysis of the secretory precursor processing machinery of *Bacillus subtilis*: identification of a eubacterial homolog of archaeal and eukaryotic signal peptidases. *Genes Dev.* 12, 2318–2331.

Walter, P. and Blobel, G. (1981). Translocation of proteins across the ER III SRP causes signal sequence and site specific arrest of chain elongation that is released by microsomal membranes. *J. Cell Biol.* 91, 557–561.

10.9 SRP взаимодействует со своим рецептором

Обзор

Doudna, J. A. and Batey, R. T. (2004). Structural insights into the signal recognition particle. *Annu. Rev. Biochem.* 73, 539–557.

Keenan, R. J., Freymann, D. M., Stroud, R. M., and Walter, P. (2001). The signal recognition particle. *Annu. Rev. Biochem.* 70, 755–775.

Исследование

Batey, R. T., Rambo, R. P., Lucast, L., Rha, B., and Doudna, J. A. (2000). Crystal structure of the ribonucleoprotein core of the signal recognition particle. *Science* 287, 1232–1239.

Halic, M., Becker, T., Pool, M. R., Spahn, C. M., Grassucci, R. A., Frank, J., and Beckmann, R. (2004). Structure of the signal recognition particle interacting with the elongation-arrested ribosome. *Nature* 427, 808–814.

Keenan, R. J., Freymann, D. M., Walter, P., and Stroud, R. M. (1998). Crystal structure of the signal sequence-binding subunit of the signal recognition particle. *Cell* 94, 181–191.

Powers, T. and Walter, P. (1995). Reciprocal stimulation of GTP hydrolysis by two directly interacting GTPases. *Science* 269, 1422–1424.

Siegel, V. and Walter, P. (1988). Each of the activities of SRP is contained within a distinct domain: analysis of biochemical mutants of SRP. *Cell* 52, 39–49.

Tajima, S., Lauffer, L., Rath, V. L., and Walter, P. (1986). The signal recognition particle receptor is a complex that contains two distinct polypeptide chains. *J. Cell Biol.* 103, 1167–1178.

Walter, P. and Blobel, G. (1981). Translocation of proteins across the ER III SRP causes signal sequence and site specific arrest of chain elongation that is released by microsomal membranes. *J. Cell Biol.* 91, 557–561.

Walter, P. and Blobel, G. (1982). Signal recognition particle contains a 7S RNA essential for protein translocation across the ER. *Nature* 299, 691–698.

Zopf, D., Bernstein, H. D., Johnson, A. E., and Walter, P. (1990). The methionine-rich domain of the 54 kd protein subunit of the signal recognition particle contains an RNA binding site and can be crosslinked to a signal sequence. *EMBO J.* 9, 4511–4517.

10.10 Транслокон формирует пору

Исследование

Crowley, K. S. (1994). Secretory proteins move through the ER membrane via an aqueous, gated pore. *Cell* 78, 461–471.

Deshaies, R. J. and Schekman, R. (1987). A yeast mutant defective at an early stage in import of secretory protein precursors into the endoplasmic reticulum. *J. Cell Biol.* 105, 633–645.

Esnault, Y., Blondel, M. O., Deshaies, R. J., Scheckman, R., and Kepes, F. (1993). The yeast SSS1 gene is essential for secretory protein translocation and encodes a conserved protein of the endoplasmic reticulum. *EMBO J.* 12, 4083–4093.

Hanein, D., Matlack, K. E., Jungnickel, B., Plath, K., Kalies, K. U., Miller, K. R., Rapoport, T. A., and Akey, C. W. (1996). Oligomeric rings of the Sec61p complex induced by ligands required for protein translocation. *Cell* 87, 721–732.

Liao, S., Lin, J., Do, H., and Johnson, A. E. (1997). Both lumenal and cytosolic gating of the aqueous ER translocon pore are regulated from inside the ribosome during membrane protein integration. *Cell* 90, 31–41.

Mothes, W., Prehn, S., and Rapoport, T. A. (1994). Systematic probing of the environment of a translocating secretory protein during translocation through the ER membrane. *EMBO J.* 13, 3973–3982.

Simon, S. M. and Blobel, G. (1991). A protein-conducting channel in the endoplasmic reticulum. *Cell* 65, 371–380.

van den Berg, B., Clemons, W. M., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C., and Rapoport, T. A. (2004). X-ray structure of a protein-conducting channel. *Nature* 427, 36–44.

10.11 Для транслокации необходимы дополнительные включения в транслокон, а иногда и храповик в ЭПР

Обзор

Rapoport, T. A., Jungnickel, B., and Kutay, U. (1996). Protein transport across the eukaryotic endoplasmic reticulum and bacterial inner membranes. *Annu. Rev. Biochem.* 65, 271–303.

Walter, P. and Lingappa, V. (1986). Mechanism of protein translocation across the endoplasmic reticulum membrane. *Annu. Rev. Cell Biol.* 2, 499–516.

Исследование

Gorlich, D. and Rapoport, T. A. (1993). Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane. *Cell* 75, 615–630.

Matlack, K. E., Misselwitz, B., Plath, K., and Rapoport, T. A. (1999). BiP acts as a molecular ratchet during posttranslational transport of prepro-alpha factor across the ER membrane. *Cell* 97, 553–564.

10.12 Для деградации белки перемещаются обратно в цитоплазму (путем обратного переноса через мембрану ЭПР)

Обзор

Johnson, A. E. and Haigh, N. G. (2000). The ER translocon and retrotranslocation: is the shift into reverse manual or automatic? *Cell* 102, 709–712.

Tsai, B., Ye, Y., and Rapoport, T. A. (2002). Retro-translocation of proteins from the endoplasmic reticulum into the cytosol. *Nat. Rev. Mol. Cell Biol.* 3, 246–255.

Исследование

Lilley, B. N. and Ploegh, H. L. (2004). A membrane protein required for dislocation of misfolded proteins from the ER. *Nature* 429, 834–840.

Wiertz, E. J. H. J. et al. (1996). Sec61-mediated transfer of a membrane protein from the endoplasmic reticulum to the proteasome for destruction. *Nature* 384, 432–438.

Wilkinson, B. M., Tyson, J. R., Reid, P. J., and Stirling, C. J. (2000). Distinct domains within yeast Sec61p involved in post-translational translocation and protein dislocation. *J. Biol. Chem.* 275, 521–529.

Ye, Y., Shibata, Y., Yun, C., Ron, D., and Rapoport, T. A. (2004). A membrane protein complex mediates retro-translocation from the ER lumen into the cytosol. *Nature* 429, 841–847.

Zhou, M. and Schekman, R. (1999). The engagement of Sec61p in the ER dislocation process. *Mol. Cell* 4, 925–934.

10.15 Как белки встраиваются в мембраны?

Обзор

Hegde, R. S. and Lingappa, V. R. (1997). Membrane protein biogenesis: regulated complexity at the endoplasmic reticulum. *Cell* 91, 575–582.

Wickner, W. T. and Lodish, H. (1985). Multiple mechanisms of protein insertion into and across membranes. *Science* 230, 400–407.

Исследование

Borel, A. C., and Simon, S. M. (1996). Biogenesis of polytopic membrane proteins: membrane segments assemble within translocation channels prior to membrane integration. *Cell* 85, 379–389.

Do, H., Falcone, D., Lin, J., Andrews, D. W., and Johnson, A. E. (1996). The cotranslational integration of membrane proteins into the phospholipid bilayer is a multistep process. *Cell* 85, 369–378.

Heinrich, S. U., Mothes, W., Brunner, J., and Rapoport, T. A. (2000). The Sec61p complex mediates the integration of a membrane protein by allowing lipid partitioning of the transmembrane domain. *Cell* 102, 233–244.

Hessa, T., Kim, H., Bihlmaier, K., Lundin, C., Boekel, J., Andersson, H., Nilsson, I., White, S. H., and von Heijne, G. (2005). Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature* 433, 377–381.

Kim, P. K., Janiak-Spens, F., Trimble, W. S., Leber, B., and Andrews, D. W. (1997). Evidence for multiple mechanisms for membrane binding and integration via carboxyl-terminal insertion sequences. *Biochemistry* 36, 8873–8882.

Liao, S., Lin, J., Do, H., and Johnson, A. E. (1997). Both lumenal and cytosolic gating of the aqueous ER translocon pore are regulated from inside the ribosome during membrane protein integration. *Cell* 90, 31–41.

Mothes, W., Heinrich, S. U., Graf, R., Nilsson, I., von Heijne, G., Brunner, J., and Rapoport, T. A. (1997). Molecular mechanism of membrane protein integration into the endoplasmic reticulum. *Cell* 89, 523–533.

van den Berg, B., Clemons, W. M., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C., and Rapoport, T. A. (2004). X-ray structure of a protein-conducting channel. *Nature* 427, 36–44.

10.16 Посттрансляционная интеграция белков в мембрану зависит от лидерных последовательностей

Обзор

Baker, K. P. and Schatz, G. (1991). Mitochondrial proteins essential for viability mediate protein import into yeast mitochondria. *Nature* 349, 205–208.

Schatz, G. and Dobberstein, B. (1996). Common principles of protein translocation across membranes. *Science* 271, 1519–1526.

Исследование

Eilers, M. and Schatz, G. (1986). Binding of a specific ligand inhibits import of a purified precursor protein into mitochondria. *Nature* 322, 228–232.

10.17 Серия лидерных участков определяет детальное местоположение белка внутри органеллы

Обзор

Cline, K. and Henry, R. (1996). Import and routing of nucleus-encoded chloroplast proteins. *Annu. Rev. Cell Dev. Biol.* 12, 1–26.

Исследование

Hartl, F. U., Ostermann, J., Guiard, B., and Neupert, W. (1987). Successive translocation into and out of the mitochondrial matrix: targeting of proteins to the intermembrane space by a bipartite signal peptide. *Cell* 51, 1027–1037.

van Loon, A. P. G. M. et al. (1986). The presequences of two imported mitochondrial proteins contain information for intracellular and intramitochondrial sorting. *Cell* 44, 801–812.

10.18 Внешняя и внутренняя мембраны митохондрий содержат разные транслоконы

Обзор

Dalbey, R. E. and Kuhn, A. (2000). Evolutionarily related insertion pathways of bacterial, mitochondrial, and thylakoid membrane proteins. *Annu. Rev. Cell Dev. Biol.* 16, 51–87.

Neupert, W. (1997). Protein import into mitochondria. *Annu. Rev. Biochem.* 66, 863–917.

Neupert, W. and Brunner, M. (2002). The protein import motor of mitochondria. *Nat. Rev. Mol. Cell Biol.* 3, 555–565.

Исследование

Leuenberger, D., Bally, N. A., Schatz, G., and Koehler, C. M. (1999). Different import pathways through the mitochondrial intermembrane space for inner membrane proteins. *EMBO J.* 18, 4816–4822.

Ostermann, J., Horwich, A. L., Neupert, W., and Hartl, F. U. (1989). Protein folding in mitochondria requires complex formation with hsp60 and ATP hydrolysis. *Nature* 341, 125–130.

10.19 Пероксисомы обладают собственной системой транслокации белков

Обзор

Purdue, P. E. and Lazarow, P. B. (2001). Peroxisome biogenesis. *Annu. Rev. Cell Dev. Biol.* 17, 701–752.

Subramani, S., Koller, A., and Snyder, W. B. (2000). Import of peroxisomal matrix and membrane proteins. *Annu. Rev. Biochem.* 69, 399–418.

Исследование

Dodt, G. and Gould, S. J. (1996). Multiple PEX genes are required for proper subcellular distribution and stability of Pex5p, the PTS1 receptor: evidence that PTS1 protein import is mediated by a cycling receptor. *J. Cell Biol.* 135, 1763–1774.

Elgersma, Y., Elgersma-Hooisma, M., Wenzel, T., McCaffery, J. M., Farquhar, M. G., and Subramani, S. (1998). A mobile PTS2 receptor for peroxisomal protein import in Pichia pastoris. *J. Cell Biol.* 140, 807–820.

Elgersma, Y., Vos, A., van den Berg, M., van Roermund, C. W., van der Sluijs, P., Distel, B., and Tabak, H. F. (1996). Analysis of the carboxyl-terminal peroxisomal targeting signal 1 in a homologous context in *S. cerevisiae*. *J. Biol. Chem.* 271, 26375–26382.

Goldfischer, S., Moore, C. L., Johnson, A. B., Spiro, A. J., Valsamis, M. P., Wisniewski, H. K., Ritch, R. H., Norton, W. T., Rapin, I., and Gartner, L. M. (1973). Peroxisomal and mitochondrial defects in the cerebro-hepato-renal syndrome. *Science* 182, 62–64.

Gould, S. J., Keller, G. A., Hosken, N., Wilkinson, J., and Subramani, S. (1989). A conserved tripeptide sorts proteins to peroxisomes. *J. Cell Biol.* 108, 1657–1664.

Matsuzono, Y., Kinoshita, N., Tamura, S., Shimozawa, N., Hamasaki, M., Ghaedi, K., Wanders, R. J., Suzuki, Y., Kondo, N., and Fujiki, Y. (1999). Human PEX19: cDNA cloning by functional complementation, mutation analysis in a patient with Zellweger syndrome, and potential role in peroxisomal membrane assembly. *Proc. Natl. Acad. Sci. USA* 96, 2116–2121.

South, S. T. and Gould, S. J. (1999). Peroxisome synthesis in the absence of preexisting peroxisomes. *J. Cell Biol.* 144, 255–266.

Walton, P. A., Hill, P. E., and Hill, S. (1995). Import of stably folded proteins into peroxisomes. *Mol. Biol. Cell* 6, 675–683.

10.21 Sec-система транспортирует белки на внутреннюю мембрану бактериальной оболочки и сквозь нее

Обзор

Lee, C. and Beckwith, J. (1986). Cotranslational and posttranslational protein translocation in prokaryotic systems. *Annu. Rev. Cell Biol.* 2, 315–336.

Oliver, D. (1985). Protein secretion in *E. coli*. *Annu. Rev. Immunol.* 39, 615–648.

Исследование

Beck, K., Wu, L. F., Brunner, J., and Muller, M. (2000). Discrimination between SRP- and SecA/SecB-dependent substrates involves selective recognition of nascent chains by SRP and trigger factor. *EMBO J.* 19, 134–143.

Brundage, L et al. (1990). The purified *E. coli* integral membrane protein SecY/E is sufficient for reconstitution of SecA-dependent precursor protein translocation. *Cell* 62, 649–657.

Collier, D. N. et al. (1988). The antifolding activity of SecB promotes the export of the *E. coli* maltose-binding protein. *Cell* 53, 273–283.

Crooke, E. et al. (1988). ProOmpA is stabilized for membrane translocation by either purified *E. coli* trigger factor or canine signal recognition particle. *Cell* 54, 1003–1011.

Valent, Q. A., Scotti, P. A., High, S., von Heijne, G., Lentzen, G., Wintermeyer, W., Oudega, B., and Luirink, J. (1998). The *E. coli* SRP and SecB targeting pathways converge at the translocon. *EMBO J.* 17, 2504–2512.

Yahr, T. L and Wickner, W. T. (2000). Evaluating the oligomeric state of SecYEG in preprotein translocase. *EMBO J.* 19, 4393–4401.

10.22 Sec-независимая система транспорта белков *E. coli*

Обзор

Dalbey, R. E. and Kuhn, A. (2000). Evolutionarily related insertion pathways of bacterial, mitochondrial, and thylakoid membrane proteins. *Annu. Rev. Cell Dev. Biol.* 16, 51–87.

Dalbey, R. E. and Robinson, C. (1999). Protein translocation into and across the bacterial plasma membrane and the plant thylakoid membrane. *Trends Biochem. Sci.* 24, 17–22.

Исследование

Beck, K., Wu, L. F., Brunner, J., and Muller, M. (2000). Discrimination between SRP- and SecA/SecB-dependent substrates involves selective recognition of nascent chains by SRP and trigger factor. *EMBO J.* 19, 134–143.

Samuelson, J. C., Chen, M., Jiang, F., Moller, I., Wiedmann, M., Kuhn, A., Phillips, G. J., and Dalbey, R. E. (2000). YidC mediates membrane protein insertion in bacteria. *Nature* 406, 637–641.

Scotti, P. A., Urbanus, M. L., Brunner, J., de Gier, J. W., von Heijne, G., van der Does, C., Driessen, A. J., Oudega, B., and Luirink, J. (2000). YidC, the *E. coli* homologue of mitochondrial Oxa1p, is a component of the Sec translocase. *EMBO J.* 19, 542–549.

Soekarjo, M., Eisenhawer, M., Kuhn, A., and Vogel, H. (1996). Thermodynamics of the membrane insertion process of the M13 procoat protein, a lipid bilayer traversing protein containing a leader sequence. *Biochemistry* 35, 1232–1241.

Транскрипция

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

11.12 Узнавание промотора зависит от консенсусных последовательностей

- Промотор определяется по наличию коротких консенсусных последовательностей в специфичных участках ДНК.
- Консенсусные последовательности промоторов состоят из пурина (в стартовой точке транскрипции), гексамера TATAAT (центр которого расположен в положении −10), и еще одного гексамера (центр которого расположен в положении −35).
- Последовательности конкретных промоторов обычно отличаются от среднестатистических консенсусных последовательностей по одной или более нуклеотидным позициям.

11.13 В результате мутаций эффективность работы промотора может быть повышена или понижена

- Ослабляющие мутации, уменьшая сходство промотора с консенсусными последовательностями, снижают его эффективность. Усиливающие мутации оказывают противоположный эффект.
- Мутации в последовательности −35, как правило, влияют на инициирующее связывание с РНК-полимеразой.
- Мутации в последовательности −10, как правило, влияют на процесс расплетания цепей ДНК, в ходе которого закрытый комплекс превращается в открытый.

11.14 РНК-полимераза связывает молекулу ДНК с одной стороны

- Консенсусные последовательности −35 и −10 являются точками контакта РНК-полимеразы и промотора.
- Точки контакта расположены на молекуле ДНК с одной стороны.

11.15 Сверхспирализация ДНК влияет на процесс транскрипции

- Отрицательная сверхспирализация повышает эффективность работы некоторых промоторов, способствуя расплетанию цепей ДНК.
- В ходе транскрипции перед ферментом возникает положительная сверхспирализация ДНК, а позади фермента — отрицательная сверхспирализация. Оба варианта сверхспирализации должны быть релаксированы с помощью гираз и топоизомераз.

11.16 Инициация на данном промоторе зависит от типа сигма-фактора

- *E.coli* имеет несколько сигма-факторов, каждый из которых обеспечивает связывание РНК-полимеразы с отдельной группой промоторов (со схожими −35 и −10 последовательностями).
- σ^{70} является основным фактором, инициирующим транскрипцию, другие сигма-факторы активируются только при определенных условиях.

11.17 Сигма-факторы взаимодействуют непосредственно с ДНК

- Изменение конформации σ^{70}, ассоциированного с минимальным ферментом, приводит к высвобождению ДНК-связывающих регионов фактора.
- σ^{70} связывает как −35, так и −10 последовательности.

11.18 Сигма-факторы организованы в каскады

- Каскад сигма-факторов формируется в том случае, если один сигма-фактор необходим для транскрипции гена, кодирующего следующий сигма-фактор.
- Ранние гены фага SP01 транскрибируются РНК-полимеразой клетки-хозяина.
- Один из ранних генов кодирует сигма-фактор, заставляющий РНК-полимеразу транскрибировать средние гены фага.
- Два средних гена кодируют субъединицы сигма-фактора, обеспечивающего транскрипцию РНК-полимеразой поздних генов.

11.19 Споруляция у бактерий контролируется сигма-факторами

- В ходе споруляции бактериальная клетка делится на материнскую часть и внутреннюю спору, высвобождающуюся в результате лизиса материнской клетки.
- При споруляции будущая спора и материнская клетка проходят через несколько стадий развития, каждый раз синтезируя новый сигма-фактор (замещающий предыдущий сигма-фактор).
- Скоординированное взаимодействие между спорой и материнской клеткой определяет время замены сигма-факторов.

11.20 Бактериальная РНК-полимераза завершает транскрипцию в определенных сайтах

- Условием окончания транскрипции является узнавание терминаторной последовательности, определяющей формирование шпилечной структуры в РНК-продукте.

11.21 Два типа терминаторов *E.coli*

- Независимые терминаторы состоят из G-C-богатой шпильки РНК и последующего за ней U-богатого региона, на котором происходит терминация.

11.22 Как работает ρ-фактор?

- ρ-фактор представляет собой терминирующий белок, который связывает насцентную цепь РНК в сайте *rut*, движется вдоль нее и отсоединяет от РНК-ДНК гибрида при достижении РНК-полимеразы.

11.23 Антитерминация — специальный регуляторный механизм

- Терминация может быть предотвращена воздействием антитерминирующего белка на РНК-полимеразу, что вызывает прочитывание ферментом специфического терминатора или терминаторов.
- Фаг лямбда имеет два антитерминирующих белка, pN и pQ, действующих на разные транскрипционные единицы.

11.24 Антитерминация опосредована сайтами, отличными от терминаторов

- Сайт, который связывает антитерминирующий белок, расположен в транскрипционной единице выше сайта терминации.
- Местоположение сайта антитерминации в разных случаях варьирует — он может располагаться в кодирующем регионе или даже в промоторе.

11.25 Факторы терминации и антитерминации взаимодействуют с РНК-полимеразой

- Для взаимодействия pN фага лямбда с РНК-полимеразой необходимо несколько бактериальных белков.
- Эти белки также вовлечены в антитерминацию в опероне *rrn* бактерии-хозяина.
- Взаимодействие антитерминирующего фактора pQ λ-фага с ДНК в области промотора может принимать различные формы.

11.26 Резюме



РИС. 11.1 Функция РНК-полимеразы заключается в копировании РНК с одной из цепей двухцепочечной ДНК

11.1 Введение

Транскрипция — это синтез цепи РНК, соответствующей одной из нитей молекулы ДНК-дуплекса. Когда мы говорим «соответствующей», мы имеем в виду, что синтезированная РНК будет иметь последовательность, *идентичную* одной из цепей ДНК, называемой **кодирующей цепью**, и будет комплементарна другой цепи ДНК, называемой **матричной цепью.** РИС. 11.1 показывает соответствие между двухцепочечной ДНК и одноцепочечным транскриптом.

Синтез РНК катализирует фермент **РНК-полимераза**. Транскрипция запускается, когда РНК-полимераза связывается со специальным регионом, расположенным вначале гена, — **промотором.** Промотор расположен рядом с первой парой оснований, транскрибирующихся в последовательность РНК, — **стартовой точкой.** С этого места РНК-полимераза движется вдоль матрицы, синтезируя РНК, до тех пор, пока не достигнет **терминаторной** (*t*) последовательности. Последовательность ДНК, заключенная между промотором и терминатором, называют **транскрипционной единицей** (транскриптоном). Как показано на РИС. 11.2, основной чертой транскрипционной единицы является то, что *она кодирует образование одной молекулы РНК.* Однако в состав транскрипционной единицы может входить более одного гена.

Регионы ДНК, предшествующие стартовой точке, называют последовательностями, расположенными **выше** (против хода транскрипции), а последовательности, лежащие после стартовой точки (внутри транскрипционной единицы), — последовательностями, расположенными **ниже** (по ходу) старта транскрипции. Принято записывать последовательности таким образом, чтобы транскрипция шла слева (от последовательностей, расположенных выше) направо (к последовательностям, расположенным ниже стартовой точки). Это соответствует изображению мРНК в направлении от 5′- к 3′-концу.

Последовательность ДНК часто записывают, показывая только кодирующую цепь (идентичную последовательности РНК). Положения оснований нумеруют в обоих направлениях от стартовой точки, которую

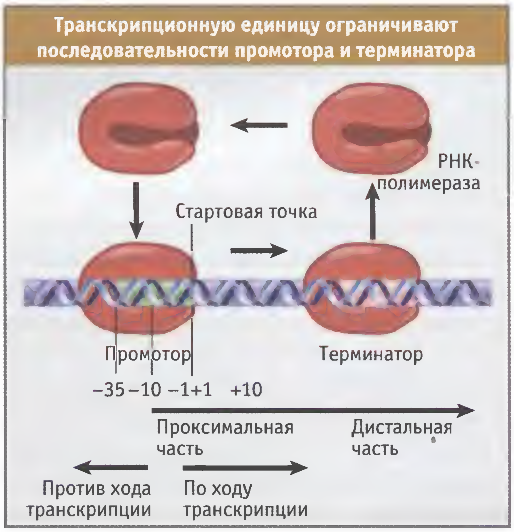


РИС. 11.2 Транскрипционная единица — последовательность ДНК, транскрибируемая в одну молекулу РНК, начинающаяся промотором и заканчивающаяся терминатором

обозначают как +1; увеличение порядковых номеров оснований происходит по ходу транскрипции. Позицию основания перед стартовой точкой обозначают как −1 и далее нумеруют отрицательными числами (основания под номером 0 не существует).

Продукт собственно транскрипции называют **первичным транскриптом**. Он представляет собой цепь РНК, состоящую из нуклеотидов, соответствующих региону между промотором и терминатором, и имеет определенные 5′- и 3′-концы. Однако первичный транскрипт почти всегда нестабилен. Например, у прокариот он либо довольно быстро деградирует (в случае мРНК), либо расщепляется с образованием зрелых продуктов (для тРНК и рРНК). У эукариот первичный транскрипт модифицируется на концах (мРНК) и/или отщепляется, давая зрелые продукты (все РНК).

Транскрипция — первая стадия механизма экспрессии генов и один из ключевых уровней, на котором

осуществляется контроль всего процесса. Множество регуляторных белков определяют, будет данный ген транскрибироваться или нет. Первый (а иногда и единственный) этап этого контроля состоит в «принятии решения» о транскрипции данного гена. Главные регуляторные события происходят на стадии инициации транскрипции, однако последующие стадии как собственно транскрипции, так и других уровней экспрессии гена тоже могут регулироваться.

Основываясь на сказанном, сформулируем два основных вопроса, касающихся экспрессии генов:

- Как РНК-полимераза находит промоторы на ДНК? Этот вопрос является частным случаем более общей формулировки: как белки находят на ДНК сайты для специфичного связывания, отличая их от всех остальных последовательностей?
- Как регуляторные белки взаимодействуют с РНК-полимеразой (и друг с другом), активируя или подавляя отдельные этапы процессов инициации, элонгации и терминации транскрипции?

В этой главе мы рассмотрим механизмы взаимодействия бактериальной РНК-полимеразы с ДНК, начиная с момента ее первичного контакта с геном, через процесс транскрипции, до высвобождения фермента после завершения синтеза транскрипта. Далее, в гл. 12 («Оперон»), описаны различные способы, с помощью которых регуляторные белки препятствуют узнаванию РНК-полимеразой отдельных генов. В гл. 13 («Регуляторная РНК») обсуждаются другие способы регуляции экспрессии генов, включая малые РНК, а также рассматривается, каким образом эти механизмы регуляции включены в общую регуляторную сеть. В гл. 14 («Стратегии фагов») мы перейдем к вопросу о том, как отдельные регуляторные взаимодействия могут быть объединены в более сложную сеть. В гл. 24 («Промоторы и энхансеры») и гл. 25 («Активация транскрипции») мы рассмотрим аналогичные процессы, происходящие в эукариотической клетке.

МОЖЕТ, УБРАТЬ СКОБКИ В НАЗВАНИЯХ ГЛАВ????????

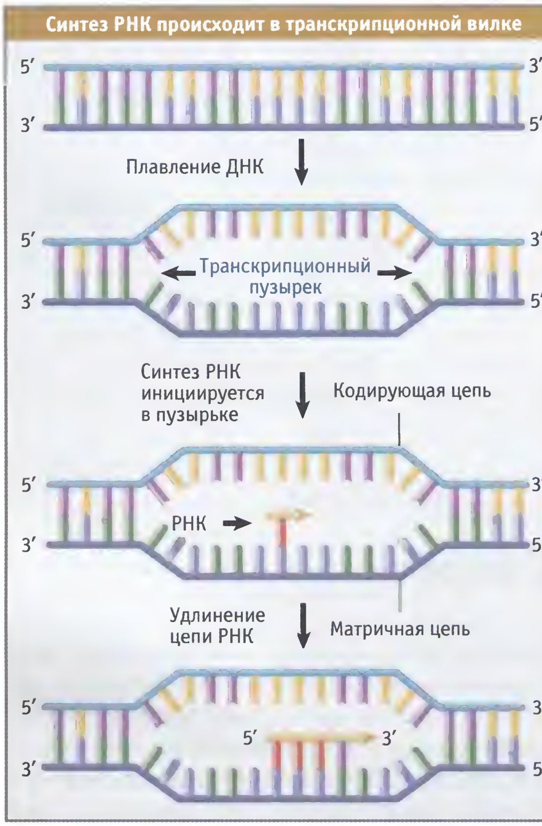


РИС. 11.3 Цепи ДНК расходятся, формируя транскрипционный пузырек. РНК синтезируется путем достраивания нуклеотидной цепи, комплементарной одной из цепей ДНК

11.2 Транскрипция происходит на коротком участке расплетенной ДНК

Ключевые понятия:

- РНК-полимераза расплетает две цепи ДНК в движущемся «пузырьке», используя одну из цепей ДНК в качестве матрицы для синтеза комплементарной последовательности РНК.
- Длина расплетенного региона (пузырька) ДНК составляет от 12 до 14 пн, длина РНК-ДНК гибрида в нем — около 8–9 пн.

В основе процесса транскрипции лежит обычное спаривание комплементарных оснований. РИС. 11.3 иллюстрирует общий принцип транскрипции. Синтез РНК происходит внутри транскрипционного пузырька, по ходу движения которого двухцепочечная молекула ДНК расплетается на отдельные цепи и одна из цепей при этом служит для матричного синтеза РНК.

Молекула РНК синтезируется в направлении от 5′-конца к 3′-концу. 3′-ОН группа последнего присоединенного нуклеотида взаимодействует с новым нуклеозид-5′-трифосфатом. Поступивший нуклеотид теряет две свои терминальные фосфатные группы (γ и β), а его α-группа используется для формирования фосфодиэфирной связи при построении цепи. Общая скорость реакции при 37 °С составляет ~40 пн/с (для бактериальной РНК-полимеразы), что примерно соответствует скорости трансляции (15 аминокислот/с), однако существенно ниже скорости репликации ДНК (800 пн/с).

Транскрипционный пузырек формируется в тот момент, когда РНК-полимераза связывается с промотором. Как показано на РИС. 11.4, с движением РНК-полимеразы вдоль ДНК перемещается и расплетенный участок, а цепь РНК при этом удлиняется. Процесс спаривания оснований и поступления в пузырек новых нуклеотидов катализируется и контролируется ферментом.

Структура пузырька внутри РНК-полимеразы детально изображена на РИС. 11.5. Двигаясь вдоль матрицы, РНК-полимераза расплетает дуплекс ДНК перед пузырьком (в точке расплетания) и восстанавливает двухцепочечную структуру ДНК позади пузырька (в точке воссоединения).

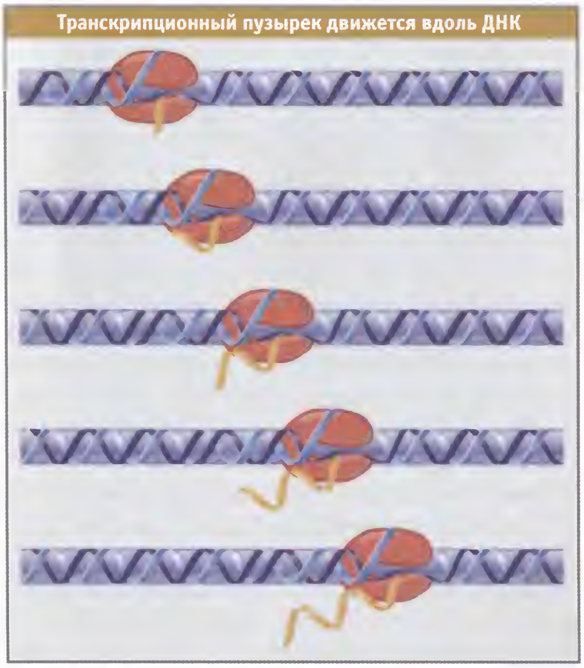


РИС. 11.4 Синтез РНК происходит в транскрипционном пузырьке, который представляет собой область локального расплетания цепей ДНК. Синтез осуществляется путем спаривания оснований с одной из цепей ДНК по принципу комплементарности. По мере движения транскрипционного пузырька ДНК восстанавливает позади него свою дуплексную структуру, высвобождая при этом синтезированную цепь РНК

Транскрипционный пузырек включает в себя ~ 12–14 пн, однако длина РНК-ДНК гибрида внутри него меньше.

Основные изменения в топологии ДНК распространяются на один виток спирали, однако непонятно, какая именно часть этого региона в действительности образует гибрид с РНК в каждый конкретный момент. РНК-ДНК дуплекс является небольшим по протяженности и короткоживущим. По мере движения фермента двухцепочечная структура ДНК восстанавливается и РНК вытесняется в виде свободной полинуклеотидной цепи. В каждый конкретный момент около 25 последних рибонуклеотидов растущей цепи находятся в комплексе с ДНК и ферментом.

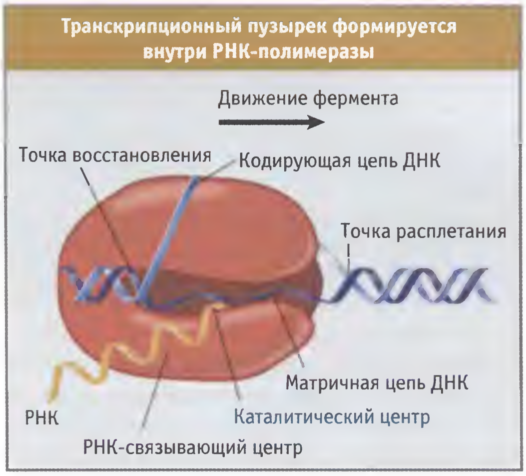


РИС. 11.5 Транскрипционный пузырек формируется внутри бактериальной РНК-полимеразы, который, двигаясь вдоль ДНК, расплетает ДНК-дуплекс перед собой, и восстанавливает двухцепочечную структуру ДНК позади

11.3 Реакция транскрипции происходит в три стадии

Ключевые понятия:

- РНК-полимераза инициирует транскрипцию после связывания с последовательностью промотора на ДНК.
- Во время элонгации транскрипционный пузырек перемещается по ДНК, синтезируя РНК цепь от 5′-конца к 3′-концу.
- При завершении транскрипции на сайте терминации РНК диссоциирует, а ДНК восстанавливает свою двухцепочечную структуру.

Как показано на **РИС. 11.6**, реакция транскрипции происходит в несколько стадий, во время которых формируется пузырек, начинается синтез РНК, пузырек движется вдоль матрицы ДНК, и, в конечном итоге, процесс терминируется:

- *Распознавание матрицы* начинается со связывания РНК-полимеразы с двухцепочечной молекулой ДНК в области промотора с образованием так называемого «закрытого комплекса». Затем цепи ДНК расплетаются и формируют «открытый комплекс», создающий условия для спаривания матрицы с рибонуклеотидами. Транскрипционный пузырек образуется в результате локального расплетания цепей ДНК, начинающегося от участка связывания ДНК с РНК-полимеразой.
- **Инициация** включает в себя синтез первых нуклеотидных звеньев цепи РНК. Фермент остается в области промотора до тех пор, пока длина синтезируемой цепи не достигнет ~9 нуклеотидов. Фаза инициации может затянуться в результате серии неудачных попыток, в ходе которых фермент синтезирует короткие транскрипты, высвобождает их и начинает синтез РНК заново. Фаза инициации заканчивается, когда ферменту удается продолжить синтез цепи и сдвинуться с промоторного участка. Промотор, таким образом, можно определить как *последовательность ДНК, необходимую для образования комплекса РНК-полимеразы и матрицы, а также обеспечивающую стадию инициации.* Неудачная (абортивная) инициация, возможно, представляет собой синтез цепи РНК, заполняющей собой активный центр фермента. Если РНК высвобождается, инициация прерывается и должна быть начата заново. Инициация заканчивается,

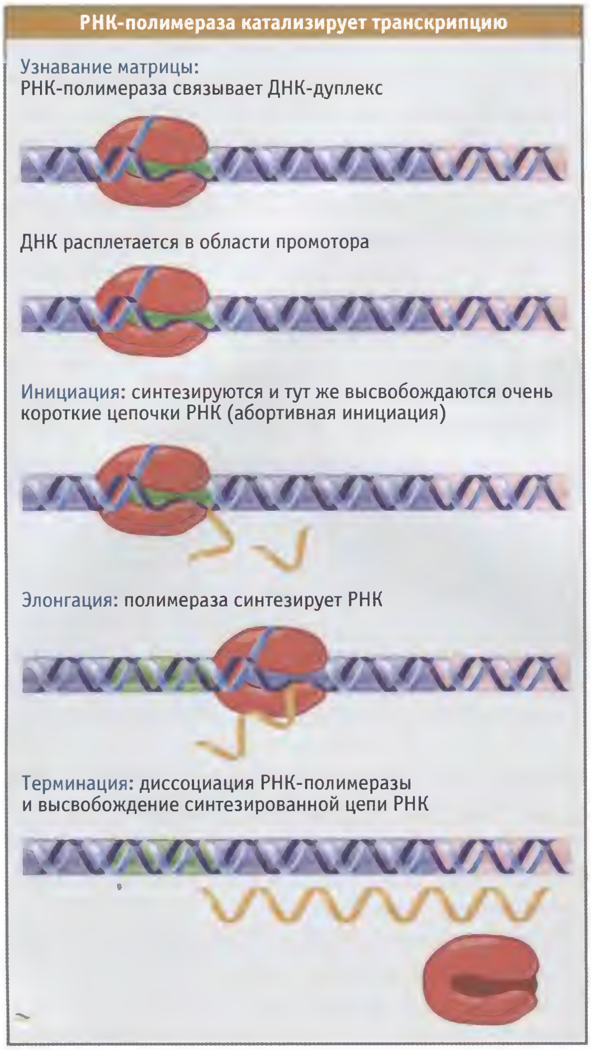


РИС. 11.6 Транскрипция происходит в четыре стадии: РНК-полимераза связывается с промотором, расплетает ДНК, во время инициации фермент остается неподвижным, во время элонгации движется вдоль матрицы ДНК и диссоциирует от нее во время терминации

когда ферменту удается продвинуться вдоль матрицы, переместив в свой активный центр следующий участок ДНК.

- В ходе **элонгации** фермент движется вдоль ДНК, обеспечивая рост цепи РНК. По мере своего движения, фермент расплетает дуплекс ДНК, постоянно формируя участок одноцепочечной матрицы. Нуклеотиды ковалентно присоединяются к 3′-концу растущей цепи РНК, образуя в месте расплетания гетеродуплекс РНК и ДНК. Позади участка расплетания матричная цепь ДНК вновь реассоциирует со второй цепью, восстанавливая двойную спираль, а РНК высвобождается в виде одноцепочечной молекулы. *Элонгация представляет собой движение транскрипционного пузырька, в ходе которого цепи ДНК расплетаются таким образом, чтобы одна из цепей могла служить матрицей для синтеза РНК.*
- **Терминация** представляет собой распознавание специальной последовательности ДНК, прекращающей дальнейший рост цепи РНК. Для терминации транскрипции необходимо, чтобы прекратилось формирование фосфодиэфирных связей, а транскрипционный комплекс диссоциировал от ДНК. Когда в цепь РНК включается последнее основание, транскрипционный пузырек исчезает, а гетеродуплекс РНК-ДНК распадается. Фермент и новосинтезированная РНК диссоциируют, а ДНК восстанавливает свою двухцепочечную структуру. *Последовательность ДНК, необходимую для обеспечения этой стадии, называют терминатором.*

Согласно общепринятой модели, элонгация представляет собой монотонный процесс, в ходе которого фермент передвигается по ДНК на одно основание после встраивания в цепь РНК каждого следующего нуклеотида. Иногда в этом механизме могут возникать отклонения, в частности в виде пауз в работе РНК-полимеразы. Одним из примеров такого явления может служить остановка «передней части» фермента, в то время как «задняя часть» продолжает двигаться, уменьшая тем самым зону покрытия ДНК. Это явление послужило источником появления «пяденичной» модели транскрипции, согласно которой фермент прерывисто с противоположных своих концов сжимает и высвобождает цепочку ДНК (как движется гусеница пяденицы). Однако возможно, что подобный механизм описывает скорее ошибочные ситуации, чем нормально протекающую транскрипцию.

11.4 РНК-полимераза фага Т7 является удобной модельной системой

Ключевые понятия:

- РНК-полимеразы фагов Т3 и Т7 представлены всего одной полипептидной цепью с «минимальной» активностью, позволяющей им узнавать лишь несколько фаговых промоторов.
- Рентгеноструктурный анализ комплекса РНК-полимеразы фага Т7 с ДНК позволяет определить ДНК-связывающий регион и активный центр фермента.

Существование у некоторых фагов небольших РНК-полимераз, представленных лишь одной полипептидной цепью, показывает, каким может быть «минимальный» транскрипционный аппарат. Эти ферменты узнают всего несколько промоторов в фаговой ДНК (и не могут узнавать остальные). Такие РНК-полимеразы являют собой простую модель, на примере которой можно охарактеризовать процесс связывания фермента с ДНК на стадии инициации транскрипции.

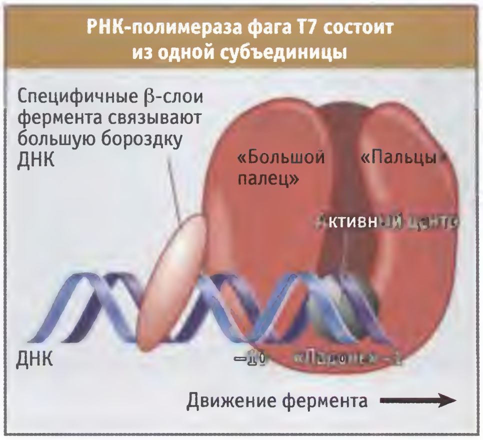


РИС. 11.7 РНК-полимераза фага T7 с помощью специфической петли связывается с промотором в положениях от −7 до −11 выше промотора, при этом основания под номерами от −1 до −4 входят в активный центр фермента

РНК-полимеразы, кодируемые родственными фагами T3 и T7, представляют собой белки, состоящие из единственной полипептидной цепи с молекулярной массой <100 кДа. При 37 °C они синтезируют РНК со скоростью ~ 200 нуклеотидов/с, что значительно превышает скорость работы бактериальной РНК-полимеразы.

РНК-полимераза фага T7 гомологична ДНК-полимеразе и имеет похожую структуру. ДНК в структуре фермента располагается в «ладони», окруженной «большим пальцем» и остальными «пальцами» (см. рис. 18.7). Рентгеноструктурный анализ РНК-полимеразы фага T7 дает нам четкое представление о том, как выглядит активный центр фермента, участвующий в процессе транскрипции.

Как показано на РИС. 11.7, РНК-полимераза фага T7 взаимодействует с большой бороздкой ДНК, связывая основания, расположенные выше стартовой точки транскрипции. Во взаимодействии фермента и ДНК участвует *петля специфичности*, представляющая собой β-слой. Эта черта является уникальной для данного класса РНК-полимераз (и не обнаружена у гомологичных ДНК-полимераз). Общим свойством всех РНК-полимераз является способность распознавать на ДНК определенные основания, расположенные перед транскрибируемой последовательностью.

Во время инициации транскрипции и добавления первых нескольких нуклеотидов конформация фермента не изменяется и матричная цепь ДНК продолжает оставаться в его активном центре в «сжатом» виде. Активный центр может удерживать транскрипт, состоящий из 6-9 нуклеотидов. Переход от инициации транскрипции к элонгации определяется моментом, в который фермент начинает свое движение вдоль ДНК. Это происходит, когда растущий транскрипт выходит за пределы активного центра и взаимодействует с петлей специфичности. РНК выходит на поверхность фермента после того, как будут синтезированы 12–14 нуклеотидов. Похожие черты характерны и для бактериальных РНК-полимераз.

11.5 На основе рентгеноструктурного анализа РНК-полимеразы можно построить модель движения фермента

Ключевые понятия:

- ДНК движется по специальной бороздке в РНК-полимеразе дрожжей, делая резкий поворот в активном центре.
- Белковый мостик меняет конформацию фермента для того, чтобы контролировать поступление нуклеотидов в активный центр.

Сейчас, благодаря наличию структур псевдокристаллов бактериальной РНК-полимеразы и РНК-полимеразы дрожжей, мы располагаем большей информацией о строении и функциях этих ферментов. Размеры РНК-полимеразы бактерий составляют ~90 × 95 × 160 Å; эукариотическая РНК-полимераза больше, но не такая вытянутая. Рентгеноструктурный анализ выявил общий тип строения этих ферментов, характеризующийся наличием на поверхности фермента канала (или бороздки) для прохождения ДНК. Пример такого канала в бактериальной РНК-полимеразе приведен на РИС. 11.8. Длина кана-



РИС. 11.8 β- (голубой) и β′-субъединицы (розовый) РНК-полимеразы формируют канал для матрицы ДНК. Синтез РНК транскрипта (оранжевый) только что начался; матричная цепь ДНК (красный) и кодирующая цепь (желтый) расплетены в транскрипционной вилке. Любезно предоставлено Seth Darst, Rockefeller University

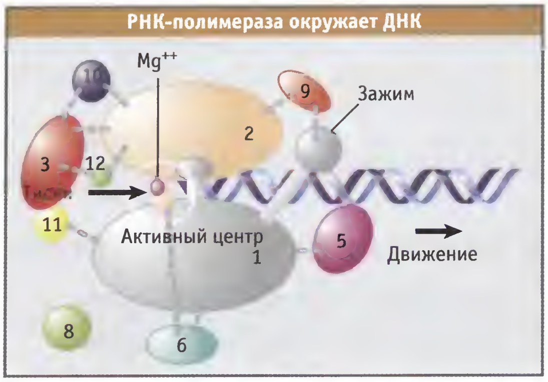


РИС. 11.9 Схематичное изображение структуры кристалла РНК-полимеразы и расположения ее 10 субъединиц. Цветовые обозначения субъединиц соответствуют обозначениям на следующих рисунках

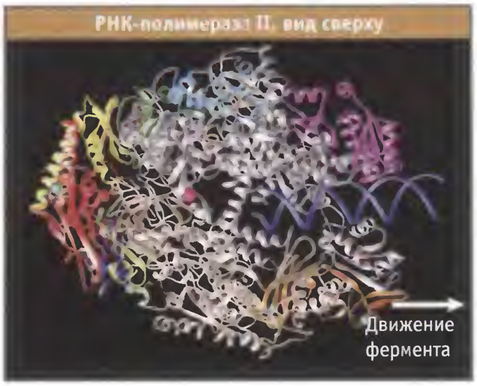


РИС. 11.10 Вид сверху структуры РНК-полимеразы II дрожжей демонстрирует, как ДНК захватывается тисками и удерживается в активном центре, содержащем ионы Mg^{2+}. Любезно предоставлено Roger Kornberg, Stanford University School of Medicine



РИС. 11.11 Вид сбоку структуры РНК-полимеразы II дрожжей демонстрирует, что ДНК окружена белком приблизительно на 270°. Любезно предоставлено Roger Kornberg, Stanford University School of Medicine

ла позволяет удерживать в нем 16 пн (в бактериальном ферменте) или ~ 25 пн (в эукариотическом), однако эта последовательность составляет лишь часть всей связанной во время транскрипции ДНК. Поверхность фермента заряжена большей частью отрицательно, а бороздку выстилают положительно заряженные аминокислоты, взаимодействующие с отрицательно заряженными фосфатными группами ДНК.

РНК-полимераза дрожжей — крупная белковая структура, состоящая из 12 субъединиц. (см. разд. 24.2 «Эукариотические РНК-полимеразы состоят из большого количества субъединиц»). 10 субъединиц дрожжевой РНК-полимеразы II расположены в структуре псевдокристалла, как показано на РИС. 11.9. Каталитический сайт формируется углублением между двумя большими субъединицами (№ 1 и № 2), которыми ДНК захватывается словно в тиски, как только попадает в РНК-полимеразу. Субъединицы 4 и 7 не обозначены на этой структуре; они формируют субкомплекс, диссоциирующий от полного фермента. По всем ключевым параметрам данная структура схожа со структурой РНК-полимеразы бактерий. Более детально это сходство можно видеть из структуры псевдокристалла, изображенной на РИС. 11.10. На РИС. 11.11 показано, что РНК-полимераза практически окружает ДНК. В активном центре фермента обнаружены ионы Mg^{2+}, необходимые для функционирования фермента. ДНК удерживается в активном центре субъединицами 1, 2 и 6. Как изображено на РИС. 11.12, для того чтобы войти в активный центр фермента, ДНК должна обогнуть прилегающую заслонку белка. Длина гетеродуплекса ограничивается другим элементом структуры, получившим название «руля». Вероятно, нуклеотиды входят в активный центр снизу, через пору в структуре белка.

Детальное рассмотрение структуры активного центра на РИС. 11.13 показывает, что транскрипционный пузырек включает в себя участок ДНК-РНК гибрида длиной 9 пн. В месте изгиба молекулы ДНК двухцепочечная спираль расплетается. По мере движения фермента вдоль ДНК, основания матричной цепи в точке изгиба молекулы ориентируются таким образом, чтобы взаимодействовать с нуклеотидами в месте их поступления в активный центр. 5′-конец РНК отделяется от ДНК после того, как наталкивается на руль фермента (см. рис. 11.12).

Несмотря на то что до начала транскрипции двойная спираль ДНК представляет собой жесткую структуру, отдельные цепи в транскрипционном пузырьке (после плавления молекулы) довольно пластичны. Это дает ДНК возможность формировать изгиб в активном центре РНК-полимеразы. Как же молекула ДНК попадает в полимеразу минуя заслонку? Для того чтобы это стало возможным, фермент должен претерпеть существенные конформационные изменения. К заслонке



РИС. 11.12 При входе в активный центр фермента ДНК должна обогнуть заслонку. Нуклеотиды, вероятно, входят в активный центр через пору в структуре белка

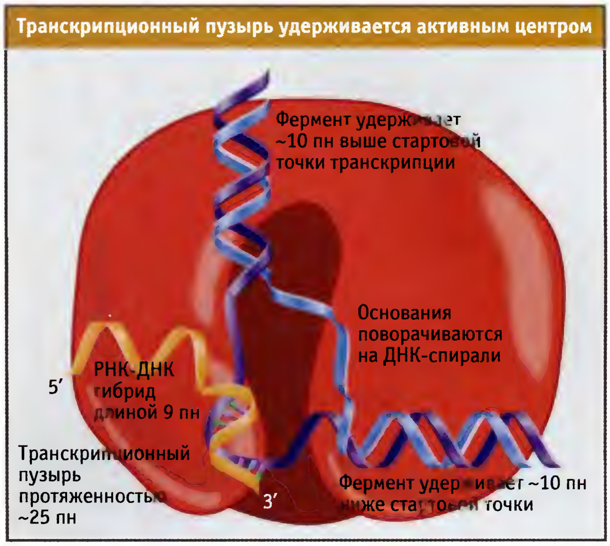


РИС. 11.13 При детальном рассмотрении видно, что ДНК при прохождении активного центра приходится делать резкий поворот

прилегает так называемый зажим. Когда РНК-полимераза находится в свободном состоянии, зажим разворачивается, так, чтобы позволить молекуле ДНК пройти внутрь фермента. После расплетания нитей ДНК и формирования транскрипционного пузырька зажим занимает по отношению к заслонке исходное положение. Любая полимераза в ходе своей работы должна, с одной стороны, образовывать тесные контакты одновременно с субстратной нуклеиновой кислотой и продуктом реакции, а с другой — должна обладать способностью разрушать и вновь восстанавливать эти контакты на каждом цикле добавления нового нуклеотида. Рассмотрим ситуацию, изображенную на РИС. 11.14. Полимераза формирует серию специфичных контактов с основаниями в определенных позициях. К примеру, контакт под номером «1» образуется с нуклеотидом, расположенным на конце растущей цепи, контакт «2» — с нуклеотидом матрицы, комплементарным следующему включаемому в цепь РНК нуклеотиду. Заметим, однако, что нуклеотиды в этих положениях меняются при добавлении к растущей цепи РНК каждого следующего нуклеотида! Верхняя и нижняя схемы, изображенные на рисунке, демонстрируют механизм добавления нуклеотида к строящейся цепи. Разница между ними лишь в том, что на нижней схеме растущая цепь удлинилась на один нуклеотид. Геометрия обоих комплексов совпадает, однако в схеме, расположенной в нижней части рисунка, контакты «1» и «2» сформированы основаниями матричной цепи, расположенными на одну позицию дальше. Как показано на центральной схеме, после того, как основание добавлено, но до того, как фермент сдвинется вдоль матричной цепи, контакты, образуемые полимеразой, должны быть разрушены и восстановлены уже с новыми основаниями, занимающими свои позиции после движения фермента.

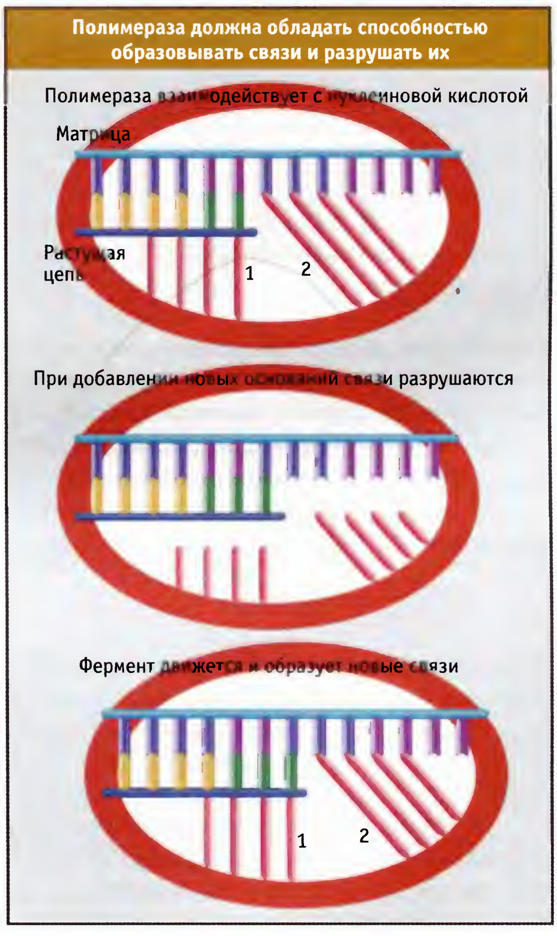


РИС. 11.14 Любая полимераза обладает способностью разрушать и вновь восстанавливать контакты с нуклеотидами в определенных положениях согласно своей структуре. Нуклеотиды в этих контактах меняются по мере движения фермента вдоль матрицы

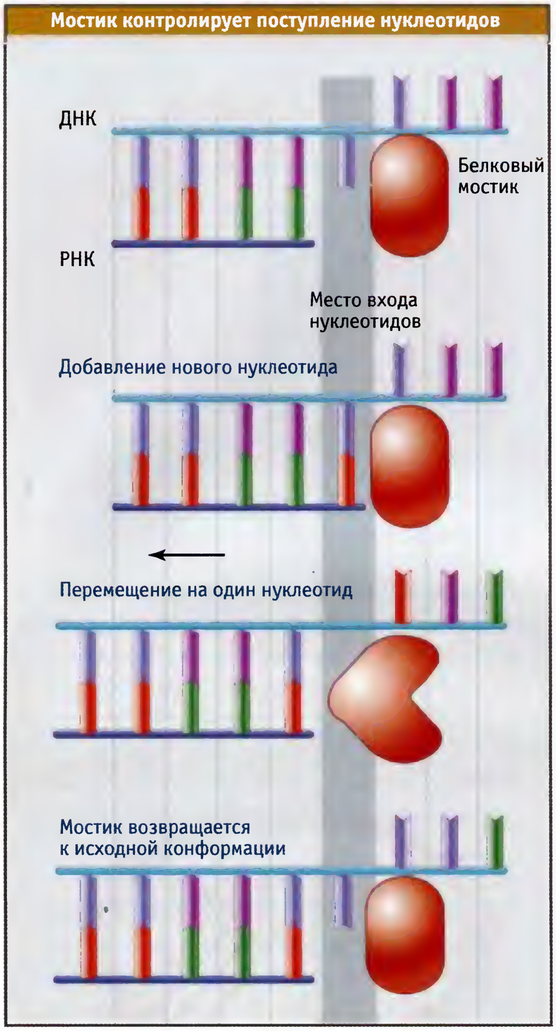


РИС. 11.15 Цикл элонгации РНК-полимеразы начинается от мостика, прилегающего к месту входа нуклеотидов. После добавления нового нуклеотида фермент сдвигается на одно основание, а мостик изменяет свою конформацию, сгибаясь для того, чтобы сохранить контакт с только что присоединенным нуклеотидом. После высвобождения мостик приобретает исходную конформацию и повторяет цикл

Структура РНК-полимеразы позволяет нам предположить, каким образом фермент сохраняет контакт с субстратом во время разрушения и восстановления связей. Существует так называемый мостик — элемент структуры белка, прилегающий к активному центру (см. рис. 11.12). Наличие мостика характерно как для бактериальных ферментов, так и для РНК-полимераз дрожжей, однако структуры псевдокристаллов того и другого показывают, что они имеют разные формы. В одном случае мостик изогнут, в другом — прямой. На РИС. 11.15 показано, как изменения конформации мостика влияют на транслокацию фермента вдоль нуклеотидной цепи.

В начале цикла транслокации мостик имеет прямую конформацию и примыкает к месту входа нуклеотидов. Мостик контактирует с вновь поступившими нуклеотидами, после чего белок сдвигается вдоль субстрата на одно основание. Мостик изменяет свою конформацию, сгибаясь для того, чтобы сохранить контакт с только что присоединенным нуклеотидом. В этой конформации мостик заслоняет собой место входа нуклеотидов. Заканчивая цикл, «мостик» приобретает исходную, прямую, конформацию, открывая доступ новым нуклеотидам. Мостик действует как храповик, освобождающий цепи ДНК и РНК для транслокации и придерживающий при этом конец растущей цепи.

11.6 Бактериальные РНК-полимеразы состоят из нескольких субъединиц

Ключевые понятия:

- Основной фермент бактериальной РНК-полимеразы представляет собой многосубъединичный комплекс с молекулярной массой ~500кДа и общей структурой $\alpha_2\beta\beta'$.
- ДНК связывается с ферментом в канале, образованном β- и β'-субъединицами.

В настоящее время лучше всего охарактеризованы РНК-полимеразы грамотрицательных бактерий, в частности фермент *E. coli*. *У эубактерий синтез мРНК, рРНК и тРНК осуществляет одна и та же РНК-полимераза.* Общее число молекул РНК-полимераз, присутствующее в клетках *E. coli*, в среднем составляет около 7000. Значительная часть из них может быть вовлечена в транскрипцию; в зависимости от условий роста в синтезе РНК может быть занято одновременно от 2000 до 5000 молекул фермента.

Полный фермент, или **холофермент**, *E. coli* имеет молекулярную массу ~ 465 кДа. Его субъединичная структура представлена на РИС. 11.16.

β- и β'-субъединицы формируют каталитический центр фермента. Их последовательности схожи с последовательностями самых больших субъединиц эукариотических РНК-полимераз (см. разд. 24.2 «Эукариотические РНК-полимеразы состоят из большого количества субъединиц»), что позволяет предположить существование некоторых общих черт в работе всех РНК-полимераз. β-субъединица перекрестно взаимодействует с матрицей ДНК, РНК-продуктом и субстратными рибонуклеотидами; мутации в гене *rpoB* влияют на все этапы транскрипции. Мутации в *rpoC* демонстрируют, что β'-субъединица также вовлечена в процесс на всех его стадиях.

α-субъединицы, наряду с β и β', входят в состав минимального фермента. У *E.coli*, инфицированной фагом T4, α-субъединица модифицируется в результате АДФ-рибозилирования аргинина. Такая модификация ведет к снижению аффинности связывания с промото-

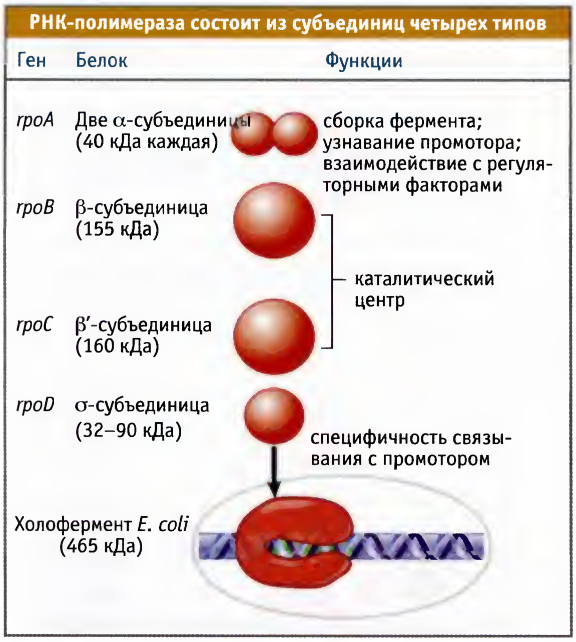


РИС. 11.16 Эубактериальные РНК-полимеразы состоят из субъединиц четырех типов; α-, β- и β′-субъединицы имеют относительно постоянный размер у разных видов бактерий, σ-субъединица гораздо более вариабельна

ром. Это позволяет предположить, что α-субъединицы играют определенную роль в узнавании промотора. Кроме того, α-субъединица участвует во взаимодействии РНК-полимеразы и некоторых факторов регуляции транскрипции.

σ-субъединица необходима исключительно для распознавания промотора, о функционировании этой субъединицы мы располагаем гораздо большей информацией, чем о любой другой субъединице (см. разд. 11.7 «РНК-полимераза состоит из минимального фермента и сигма-фактора»).

Из структуры псевдокристалла бактериального фермента (см. рис. 11.8) видно, что канал для ДНК расположен между β- и β′-субъединицами (α-субъединицы на этом рисунке не видны). В активном центре, где происходит синтез цепи РНК, ДНК находится в расплетенном состоянии. С помощью ряда экспериментов были определены участки, в которых субъединицы РНК-полимеразы контактируют с ДНК. Они показаны на РИС. 11.17. β- и β′-субъединицы взаимодействуют с ДНК в нескольких точках, расположенных ниже активного центра. Несколько контактов с кодирующей цепью они формируют в районе транскрипционной вилки, стабилизируя таким образом разделенные цепи ДНК. РНК контактирует с ферментом в основном в области транскрипционного пузырька.

Рифампицин (член семейства антибиотиков группы рифамицинов) блокирует работу бактериальной РНК-полимеразы. Он является главным лекарственным препаратом, используемым при лечении туберкулеза. Структура псевдокристалла РНК-полимеразы в комплексе с рифампицином объясняет механизм действия антибиотика: рифампицин размещается в кармане, образуемом β-субъединицей. Несмотря на то что место связывания расположено на расстоянии свыше 12 Å от активного центра, рифампицин блокирует элонгацию РНК. Не позволяя цепи РНК удлиниться более чем на 2–3 нуклеотида, антибиотик ингибирует транскрипцию.

РНК-полимераза, которой первоначально приписывалась способность только наращивать цепь РНК на матрице ДНК, сейчас рассматривается как компонент более сложного аппарата, вовлеченного в процесс транскрипции. *Способность катализировать синтез РНК является «минимальной» функцией РНК-полимеразы.* Фермент контролирует спаривание субстратных рибонуклеотидов с ДНК и катализирует образование фосфодиэфирных связей между ними.

Все субъединицы минимального фермента, участвующие в элонгации, необходимы также и для процессов инициации и терминации. Транскрипционные комплексы, однако, различаются по набору дополнительных полипептидов, участие которых необходимо для инициации и терминации транскрипции определенных генов. Очевидна аналогия с разделением труда между рибосомой и факторами белкового синтеза при трансляции.

РНК-полимераза *E.coli* способна транскрибировать любую из более чем 1000 транскрипционных единиц этой бактерии. В связи с этим становится необходимым существование большого числа регуляторов, способных влиять на транскрипционную активность фермента. Сложность структуры фермента отражает, скорее, его способность взаимодействовать с различными регуляторными факторами, чем какие-то особенности его каталитической активности.



РИС. 11.17 И матричная, и кодирующая цепи ДНК контактируют с β- и β′-субъединицами главным образом в области транскрипционного пузыря и ниже по ходу транскрипции. РНК также контактирует с ферментом в основном в области транскрипционного пузыря. На участках ниже пузыря (по ходу транскрипции) РНК контактов не образует, за исключением особых случаев, в которых фермент движется в обратном направлении

11.7 РНК-полимераза состоит из минимального фермента и сигма-фактора

Ключевые понятия:

- Бактериальная РНК-полимераза может быть разделена на минимальный фермент $\alpha_2\beta\beta'$, катализирующий собственно реакцию транскрипции, и сигма-субъединицу, необходимую только для инициации процесса.
- Сигма-фактор изменяет характер взаимодействия РНК-полимеразы с ДНК таким образом, что аффинность связывания фермента со случайным участком ДНК понижается, а с промотором — возрастает.
- Константа связывания РНК-полимеразы с разными промоторами варьирует в диапазоне шести порядков, определяя частоту, с которой транскрипция инициируется на каждом из этих промоторов.

Холофермент ($\alpha_2\beta\beta'\sigma$) можно разделить на два компонента: **минимальный фермент** (core enzyme) ($\alpha_2\beta\beta'$) и **сигма-фактор** (σ-полипептид). *Только полный холофермент может инициировать транскрипцию. Сигма-фактор обеспечивает образование стабильной связи между бактериальной РНК-полимеразой и ДНК в области промотора.* После того как длина цепи РНК достигнет 8–9 оснований, сигма-фактор высвобождается, и элонгация осуществляется только минимальным ферментом. *Минимальный фермент способен синтезировать цепь РНК на ДНК-матрице, но не может инициировать транскрипцию в нужном участке.*

Минимальный фермент сам по себе обладает сродством к ДНК, в основе которого лежит электростатическое взаимодействие между положительно заряженным белком и отрицательно заряженной нуклеиновой кислотой. Любую (случайную) последовательность ДНК, связывающуюся с базовой РНК-полимеразой, называют **свободным участком связывания.** ДНК при таком взаимодействии не претерпевает никаких изменений, сохраняя двухцепочечную структуру. При этом такой комплекс достаточно стабилен, период его полураспада на ДНК и минимальный фермент составляет ~60 мин. *Однако минимальный фермент не отличает промоторы от других последовательностей ДНК.*

Как показано на , сигма-фактор принципиально изменяет характер взаимодействия РНК-полимеразы с ДНК. *Сродство холофермента к свободным участкам связывания, а значит, и способность связываться с любыми последовательностями ДНК резко снижаются.* Константа связывания для этой реакции благодаря сигма-фактору снижается в 10^4 раз, а период полураспада такого комплекса не превышает 1 с. Таким образом, сигма-фактор существенно снижает способность фермента связываться со случайными участками ДНК.

Следует отметить, что *сигма-фактор придает ферменту способность узнавать определенные участки связывания.* Холофермент очень прочно связывается с промоторами, константа связывания для этой реакции по сравнению с базовым ферментом возрастает (в среднем) в 1000 раз, а период полураспада такого комплекса составляет несколько часов.

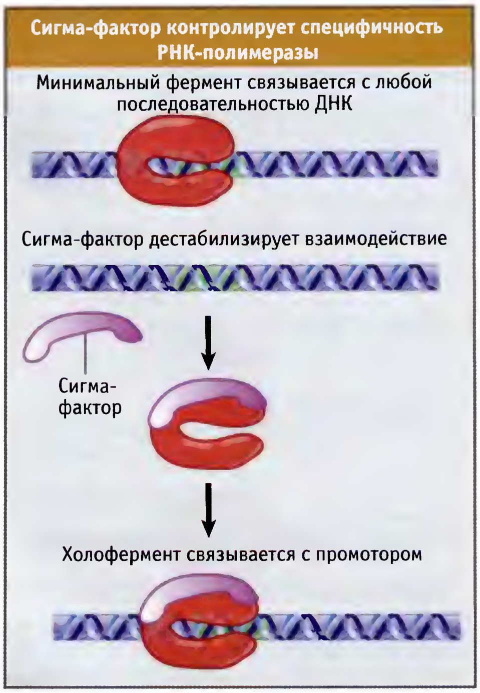


РИС. 11.18 Минимальный фермент неизбирательно связывается с любыми последовательностями ДНК. Сигма-фактор снижает сродство минимального фермента к случайным участкам связывания и сообщает ему специфичность к промоторным последовательностям

Специфичность холофермента по отношению к промоторам (по сравнению с другими последовательностями) выше в ~10^7 раз. Данная величина является усредненной, поскольку существуют вариации в скорости связывания холофермента с различными промоторами. Сродство холофермента к промотору имеет принципиальное значение для эффективности инициации с данного конкретного промотора. Константа связывания варьирует от ~10^{12} до ~10^6. Другие факторы также оказывают влияние на частоту инициации, которая в итоге варьирует от ~1 в секунду (для генов рРНК) до ~1 в полчаса (для промотора *lacI*).

11.8 После инициации транскрипции характер взаимодействия минимального фермента с сигма-фактором меняется

Ключевые понятия:

- После ассоциации РНК-полимеразы с промотором цепи ДНК расплетаются, а фермент синтезирует молекулу РНК длиной до девяти нуклеотидов.
- Удачному продолжению синтеза РНК может предшествовать серия абортивных инициаций.
- Как только растущая цепь РНК достигает в длину восьми-девяти оснований, сигма-фактор диссоциирует от РНК-полимеразы.

Сейчас мы можем рассматривать все стадии транскрипции в контексте взаимодействия между различными формами РНК-полимеразы с матрицей ДНК. Реакция инициации может быть описана при помощи параметров, перечисленных на РИС. 11.19:

- Взаимодействие холофермента с промотором начинается с формирования двойного закрытого комплекса. Термин «закрытый» означает, что ДНК в этом комплексе сохраняет свою двухцепочечную структуру. Формирование двойного закрытого комплекса обратимо, поэтому для его описания используют обычно константу равновесия (K_B). Для разных закрытых комплексов существуют значительные вариации в значениях константы равновесия.
- Закрытый комплекс переходит в **открытый** в результате плавления небольшого участка ДНК в пределах последовательности, связанной с ферментом. Ряд последовательных событий, в результате которых формируется открытый комплекс, называют **прочным связыванием.** Для промоторов, прочность связывания с которыми высока, переход к открытому комплексу необратим; такая реакция описывается константой скорости (k_2). Переход происходит быстро, при этом сигма-фактор участвует в расплетании цепей ДНК (см. разд. 11.16 «Замена сигма-фактора контролирует инициацию»).
- Следующим шагом является включение первых двух нуклеотидов и формирование фосфодиэфирной связи между ними. Это приводит к образованию **тройного комплекса** между ферментом, ДНК и синтезируемой РНК. Образование тройного комплекса описывается с помощью константы скорости k_i, которая меньше, чем константа k_2. Пока длина цепи РНК не превысит 9 оснований, присоединение нуклеотидов не требует движения фермента. После добавления каждого основания существует вероятность того, что фермент высвободит синтезируемой цепь. Это явление носит название **абортивной инициации**. В этом случае фермент должен начать синтез заново, с первого основания. Как правило, в результате ряда абортивных инициаций образуется серия очень коротких олигонуклеотидов.
- После завершения инициации сигма-фактор высвобождается, и дальнейшая элонгация осуществляется тройным комплексом, в состав которого входят кор-фермент, ДНК и синтезируемая РНК. Критическим параметром в этом процессе является достаточность скорости *ухода полимеразы с промотора для того, чтобы следующий фермент мог узнать его.* Этот параметр называют скоростью освобождения промотора; его минимальное значение, составляющее примерно 1–2 секунды, определяет максимальную частоту инициации, составляющую около 1 события в секунду. Фермент начинает движение вдоль ДНК-матрицы при достижении цепью РНК длины в 10 оснований.

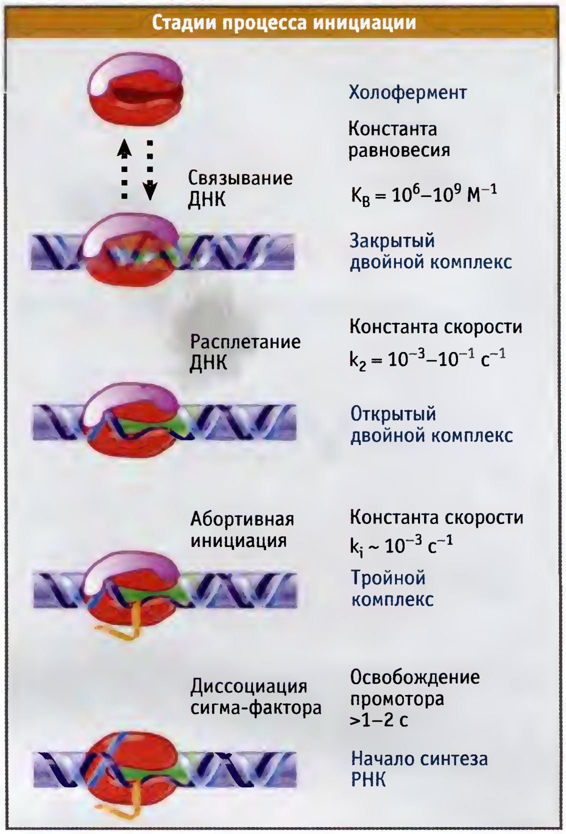


РИС. 11.19 До начала элонгации РНК-полимераза проходит через несколько стадий процесса инициации. Закрытый двойной комплекс превращается в открытый, а затем — в тройной комплекс

При связывании РНК-полимеразы с субстратом длинная ось фермента еще больше растягивается вдоль ДНК, однако во время транскрипции происходит ряд интересных конформационных изменений. Эти изменения формы и размеров описывают три варианта тройного комплекса, изображенные на РИС. 11.20:

- При первичном связывании холофермента с ДНК РНК-полимераза экранирует от 75 до 80 пн, «растягиваясь» от –55 до +20 позиции (по направлению своей длинной оси РНК-полимераза (160 Å) может покрывать только ~50 пн ДНК, что означает, что связывание фермента с бо́льшим участком ДНК является следствием некоего изгибания молекулы нуклеиновой кислоты).
- Форма РНК-полимеразы менятся при переходе от стадии инициации к стадии элонгации. Этот процесс сопровождается потерей контактов в участке между позициями –55 и –35, с ферментом остаются связаны только ~60 пн ДНК. Это согласуется с идеей о том, что часть промотора, расположенная выше по ходу транскрипции, вовлечена в первичное распознавание РНК-полимеразой, но не участвует в более поздних стадиях инициации (см. разд. 11.13 «В результате мутаций эффективность работы промотора может быть повышена или понижена »).

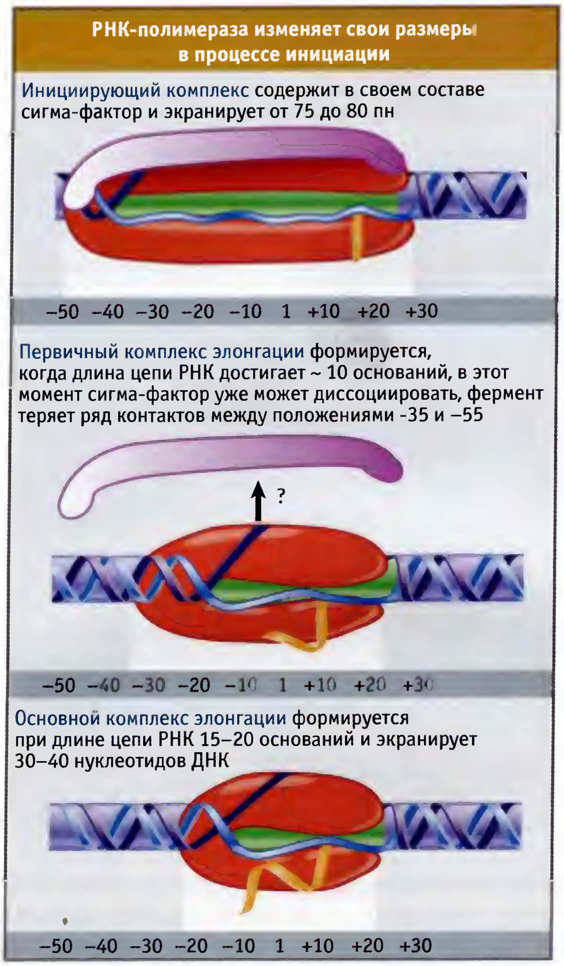


РИС. 11.20 Первоначально РНК-полимераза взаимодействует с регионом, расположенным между –55 и +20 позициями. После диссоциации сигма-фактора участок ДНК, экранируемый ферментом, уменьшается до 30 пн; в основном комплексе элонгации фермент приобретает еще более компактную структуру

- Когда длина цепи РНК увеличивается до 15–20 оснований, фермент претерпевает очередные изменения, в результате которых формируется комплекс, осуществляющий элонгацию (основной комплекс элонгации); в этой конформации фермент покрывает от 30 до 40 оснований (в зависимости от стадии цикла элонгации).

Вскоре после открытия механизм высвобождения сигма-фактора из комплекса с полимеразой стал восприниматься как догма транскрипции. Однако позже выяснилось, что это не совсем верно. Анализ элонгирующих комплексов РНК-полимеразы показал, что в 70% случаев сигма-фактор сохраняет ассоциацию с базовым ферментом. Вместе с тем треть элонгирующих РНК-полимераз не содержит сигма-фактор, в связи с чем утверждение о том, что сигма-фактор не является необходимым участником процесса элонгации, безусловно верно. В тех случаях, когда сигма-фактор остается ассоциированным с ферментом, меняется природа этой ассоциации (см. разд. 11.11 «Сигма-фактор контролирует связывание РНК-полимеразы с ДНК»).

11.9 Остановившаяся РНК-полимераза способна продолжить транскрипцию после «перезагрузки»

Ключевые понятия:

- РНК-полимераза, по какой-то причине прервавшая свою работу, может возобновить транскрипцию путем расщепления транскрипта с образованием нового 3′-конца.

РНК-полимераза должна уметь справляться с ситуациями, когда транскрипция по какой-то причине оказывается заблокирована. Это может произойти, к примеру, в случае повреждения ДНК. Хорошей моделью такой ситуации является арест элонгации *in vitro*, возникающий в результате отсутствия в смеси одного из необходимых трифосфатов. После добавления недостающего нуклеотида, для продолжения процесса фермент отщепляет 3′-концевой нуклеотид, формируя новый 3′-конец цепи РНК. Кроме РНК-полимеразы в этот процесс вовлечены дополнительные факторы. В случае с РНК-полимеразой *E. coli* возобновление работы фермента обеспечивают белки GreA и GreB. В эукариотических клетках для отщепления РНК-полимеразой II нескольких рибонуклеотидов с 3′-конца синтезируюемой цепи необходимо участие дополнительного фактора ($TF_{II}S$). На самом деле во всех случаях удаление нуклеотидов осуществляет сам каталитический центр фермента, а роль факторов GreB и $TF_{II}S$ состоит в том, чтобы превратить каталитический центр полимеразы в рибонуклеазный центр. Хотя аминокислотные последовательности этих белков не похожи, структуры псевдокристаллов их комплексов с соответствующими РНК-полимеразами позволяют предположить, что действуют они по одному механизму. И тот, и другой белок встраивают небольшой домен (в одном случае это цинковая лента, в другом — суперспираль) глубоко в структуру полимеразы, где он оказывается частью каталитического центра фермента. В результате такого встраивания две отрицательно заряженные аминокислоты оказываются расположенными вблизи первого каталитического иона магния в активном центре фермента; это обеспечивает включение в процесс второго иона магния, что и превращает каталитический центр фермента в рибонуклеазный центр.

Причиной, по которой природе понадобилось создавать описанный выше механизм, может быть периодический выход 3′-конца синтезируемого транскрипта за пределы активного центра (за счет ошибок позиционирования фермента на матрице). Отщепление 3′-концевых нуклеотидов и обратное движение фермента возвращают конец растущей цепи в активный центр, восстанавливая условия, необходимые для дальнейшего добавления нуклеотидов.

Таким образом, можно перечислить следующие активности РНК-полимеразы: расплетение и восста-

новление двухцепочечной структуры ДНК, удержание в правильной ориентации расплетенных цепей ДНК и РНК-продукта, катализ добавления нуклеотидов к растущей рибонуклеиновой цепи, отщепление 3′-концевых нуклеотидов в случае ошибки транскрипции и возобновление синтеза РНК (при участии дополнительных факторов).

11.10 Как РНК-полимераза находит последовательность промотора?

Ключевые понятия:

- Скорость, с которой РНК-полимераза находит промотор, слишком высока, чтобы быть результатом беспорядочной диффузии.
- Возможно, РНК-полимераза связывается со случайными участками ДНК и быстро перемещается по ней, пока не обнаружит промотор.

Как молекулы РНК-полимеразы распределены в клетке? На РИС. 11.21 представлена (не бесспорная) гипотеза, характеризующая распределение функционально активных молекул фермента в клетке:

- Избыток кор-фермента существует в виде закрытых неактивных комплексов, которые РНК-полимераза формирует со случайными последовательностями ДНК. Эти комплексы образуются очень быстро, а диссоциируют медленно. В свободной форме существует лишь очень небольшая часть кор-фермента (если существует вообще).
- Сигма-фактора хватает примерно на треть существующих в клетке базовых РНК-полимераз, в результате чего образуется холофермент, распределенный между случайными регионами ДНК (в неактивном состоянии) и двойными комплексами в регионах промоторов (большей частью закрытыми).
- Примерно половина имеющихся в клетке РНК-полимераз существует в виде минимального фермента, в данный момент осуществляющего транскрипцию.
- Возможно, какая-то часть холофермента не задействована. Не располагая точными данными, мы, однако, предполагаем, что эта часть очень незначительна.

Полимераза должна уметь распознавать промоторы среди других последовательностей генома. Предположительно промотор включает в себя порядка 60 нуклеотидов. Как обнаружить эту последовательность среди 4×10^6 пар оснований, составляющих полный геном *E. coli*? Следующие три рисунка иллюстрируют несколько возможных объяснений.

На РИС. 11.22 приведена простейшая модель связывания РНК-полимеразы с промотором. Эта модель предполагает случайное движение фермента по ДНК. Холофермент очень быстро ассоциирует со случайными сайтами связывания и так же быстро диссоциирует от них. Таким образом, продвигаясь вдоль молекулы ДНК, холофермент образует и разрушает ряд закрытых комплексов до тех пор, пока в процессе поиска (случайно) не натолкнется на промотор. Тогда, распознав специфическую последовательность, он может прочно связаться с ДНК и образовать открытый комплекс.

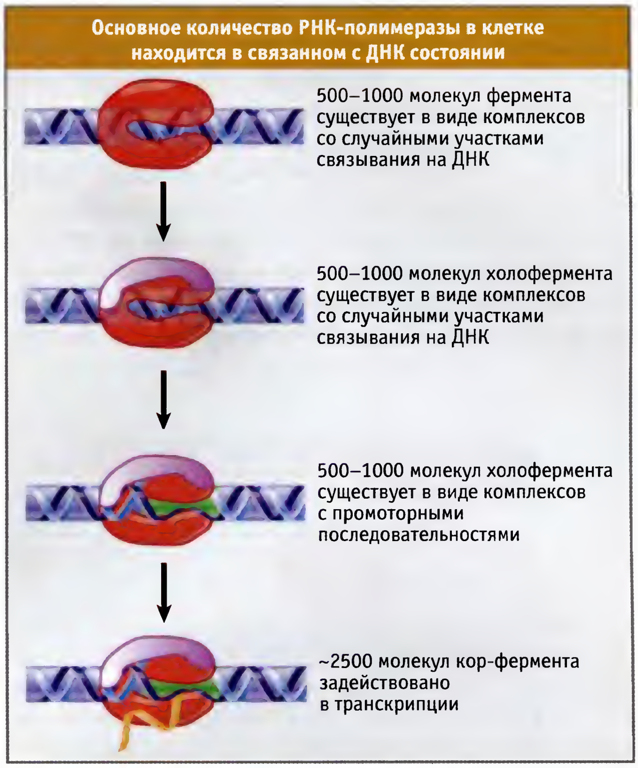


РИС. 11.21 Большая часть кор-фермента и холофермента существуют в комплексе с ДНК. В свободной форме находится лишь небольшая часть РНК-полимеразы

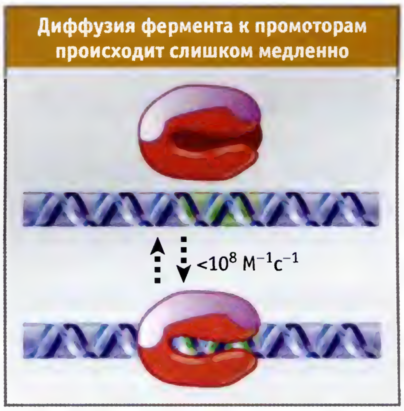


РИС. 11.22 Константа скорости связывания РНК-полимеразы с промотором выше таковой для случайной диффузии

Для того чтобы перейти от одного сайта связывания на ДНК к другому, РНК-полимераза должна диссоциировать от первого сайта, найти второй и связаться с ним. Переход от одного сайта к другому ограничивается скоростью диффузии фермента во внутриклеточной среде.

Этот показатель устанавливает верхний предел константы скорости ассоциации фермента с 60-нуклеотидной последовательностью промотора, который составляет $<10^8$ $М^{-1}$ $с^{-1}$. Однако на практике эта константа для некоторых промоторов *in vitro* составляет $\sim 10^8$ $М^{-1}$ $с^{-1}$ или даже превышает порог, определяемый скоростью диффузии. Если бы такие значения действовали *in vivo,* время, требуемое для цикла случайного связывания с неспецифической последовательностью и диссоциации от нее, было бы слишком велико для того, чтобы считать такой способ поиска промоторов возможным.

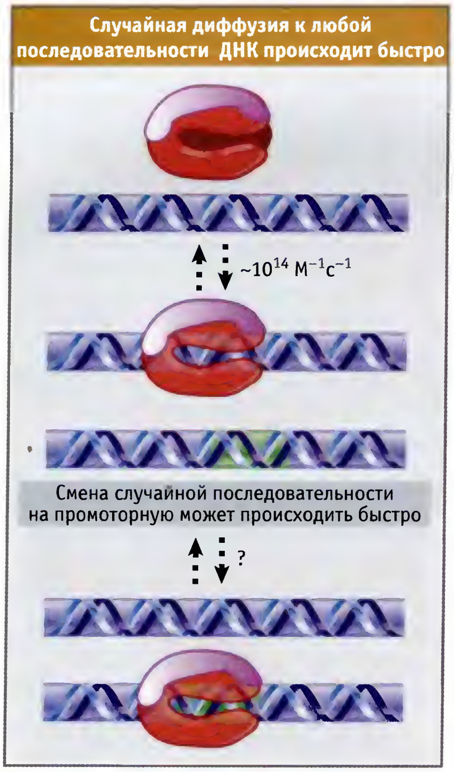


РИС. 11.23 Холофермент очень быстро связывается со случайными последовательностями ДНК, пока в результате прямой замены одной последовательности на другую не натолкнется на промотор



РИС. 11.24 РНК-полимераза не скользит вдоль ДНК

Это означает, что РНК-полимераза использует другие средства для поиска участков связывания. Как показано на РИС. 11.23, скорость реакции увеличивается за счет того, что мишенью для первичного связывания РНК-полимеразы может служить любая последовательность генома, а не только последовательность промотора. С увеличением размеров мишени для связывания увеличивается и величина константы, которая в этом случае больше не является лимитирующим фактором.

Если эта гипотеза верна, свободная РНК-полимераза связывается с ДНК и остается в контакте с ней надолго. Каким же образом фермент перемещается от случайного участка связывания к промотору? Вероятнее всего, происходит прямая замена одной связывающей последовательности на другую. Связываясь с одной последовательностью ДНК, фермент быстро меняет ее на другую и продолжает обменивать последовательности до тех пор, пока не доберется до промотора. После этого фермент образует стабильный открытый комплекс, приводящий к инициации. Процесс поиска, таким образом, значительно ускоряется, поскольку связывание и диссоциация происходят практически одновременно, и время на переключение между комплексами не тратится. Такое прямое перемещение получило название «прямой прогулки», в ходе которой фермент передвигается от участков слабого связывания к участкам прочного связывания.

Согласно еще одной гипотезе, фермент скользит вдоль ДНК посредством однонаправленного движения, изображенного на РИС. 11.24, останавливаясь только в том случае, если встретится с промотором. Однако свидетельств того, что РНК-полимераза (или другие ДНК-связывающие белки) могут функционировать по этому механизму, не обнаружено.

11.11 Сигма-фактор контролирует связывание РНК-полимеразы с ДНК

Ключевые понятия:

- Взаимодействие холофермента с сигма-фактором изменяет холофермент таким образом, что делает возможным движение минимального комплекса вдоль молекулы ДНК.

Как же решается проблема различия параметров взаимодействия фермента с матрицей в процессе инициации и в процессе элонгации? Для инициации требуется прочное взаимодействие полимеразы с конкретным участком (промотором), тогда как при элонгации необходимо прочное связывание со всеми регионами ДНК, вдоль которых происходит движение фермента. На РИС. 11.25 изображено, как эта дилемма решается путем обратимой ассоциации сигма-фактора с корферментом. Как было отмечено ранее (см. разд. 11.8

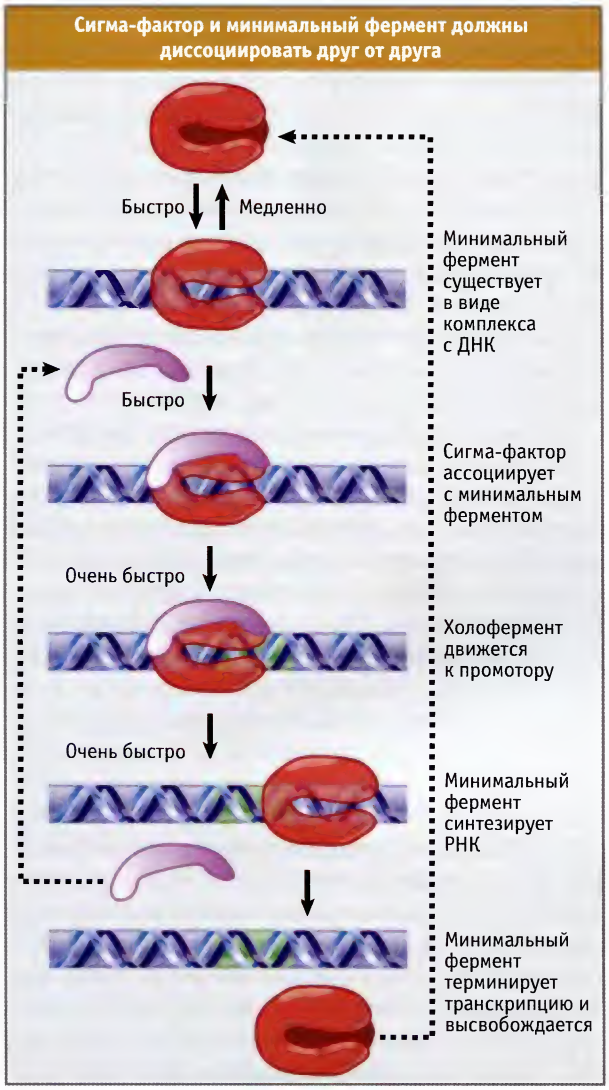


РИС. 11.25 **Взаимодействие сигма-фактора и минимального (кор-) фермента носит циклический характер**

«После инициации транскрипции характер взаимодействия минимального фермента с сигма-фактором меняется»), после инициации сигма-фактор высвобождается из комплекса с минимальным ферментом (или меняет характер связи с ним) таким образом, чтобы не участвовать далее во взаимодействии с ДНК. Поскольку число молекул сигма-фактора существенно меньше числа молекул минимального фермента, для функционирования РНК-полимеразы необходимо, чтобы сигма-фактор работал циклически. Как показано на рисунке, сразу после инициации сигма-фактор диссоциирует от минимального фермента примерно в трети случаев; вероятно, в остальных случаях диссоциация происходит позднее.

Независимо от времени, требуемого для диссоциации сигма-фактора от минимального фермента, фактор участвует только в инициации. Когда серия абортивных инициаций прекращается в результате успешной инициации синтеза РНК, фактор становится ненужным. Неясно, являются ли изменения в структуре фермента следствием прекращения серии абортивных инициаций, либо, наоборот, влекут ли изменения фермента за собой окончание неудачных попыток и начало элонгации.

Освободившись от одного минимального фермента, сигма-фактор тут же становится доступен для взаимодействия с другим. Независимо от того, диссоциирует сигма-фактор от минимального фермента или же сохраняет с ним связь (меняя характер взаимодействия на более слабый), минимальный фермент в тройном комплексе остается связан с ДНК очень прочно. Фермент как бы «заперт» в тройном комплексе до тех пор, пока не закончится элонгация. Когда же транскрипция завершается, минимальный фермент отделяется от ДНК. Незадействованный в транскрипции фермент «хранится» в виде комплекса с ДНК, связываясь со случайными ее участками. Если после инициации сигма-фактор диссоциировал от минимального фермента, то освободившийся по окончании синтеза фермент должен найти новый сигма-фактор и включиться в следующий цикл транскрипции.

Минимальный фермент сам по себе обладает высоким сродством к ДНК, которое еще увеличивается в присутствии новосинтезированной РНК. Однако прочность его связывания со случайными участками ДНК слишком высока, что не позволяет ферменту эффективно находить промоторы. Сигма-фактор значительно ускоряет процесс поиска промоторов, уменьшая стабильность случайных комплексов фермента и ДНК; в то же время, стабилизируя ассоциацию в участках прочного связывания, сигма-фактор необратимо сдвигает реакцию в сторону образования открытых комплексов. Освобождаясь от сигма-фактора (или изменяя характер связи с ним), фермент снова приобретает способность связываться с любой последовательностью ДНК, что позволяет ему продолжать транскрипцию.

Что определяет способность холофермента специфично связываться с промоторами? Сигма-фактор имеет домены, распознающие последовательность промотора. Самостоятельно сигма-фактор не связывается с ДНК, однако при образовании комплекса холофермента с промотором, сигма-фактор контактирует с ДНК в области, расположенной перед стартовой точкой транскрипции. Связывание сигма-фактора с минимальным ферментом изменяет конформацию фактора так, что он приобретает способность взаимодействовать с ДНК. N-концевой регион свободного сигма-фактора подавляет активность его ДНК-связывающего домена; после связи с минимальным ферментом это влияние исчезает (см. рис. 11.35 в разд. 11.17 «Сигма-факторы взаимодействуют непосредственно с ДНК»). Подобный механизм взаимодействия сигма-субъединицы с ДНК имеет важное значение: связывание свободного сигма-фактора с промоторами, возможно, блокировало бы инициацию транскрипции холоферментом.

11.12 Узнавание промотора зависит от консенсусных последовательностей

Ключевые понятия:

- Промотор определяется по наличию коротких консенсусных последовательностей в специфичных участках ДНК.
- Консенсусные последовательности промоторов состоят из пурина (в стартовой точке транскрипции), гексамера TATAAT (центр которого расположен в положении −10), и еще одного гексамера (центр которого расположен в положении −35).
- Последовательности конкретных промоторов обычно отличаются от среднестатистических консенсусных последовательностей по одной или более нуклеотидным позициям.

Поскольку главная функция промотора — *быть узнаваемым белками* (полимеразами или факторами регуляции транскрипции), его последовательность принципиально отличается от последовательностей, предназначенных для транскрипции и трансляции. Функциональная информация промотора исходит непосредственно из его последовательности: структура промотора представляет собой сигнал. Это классический пример так называемого *цис*-активного сайта, показанного в предыдущих главах на рис. 2.16 и 2.17. Кодирующие же области, напротив, приобретают свой смысл только после того, как закодированная в них информация будет представлена в виде РНК или белка.

Ключевым вопросом при рассмотрении взаимодействия РНК-полимеразы с промотором остается вопрос о том, как белок узнает специфическую промоторную последовательность. Имеет ли фермент активный центр, способный различать химическую структуру отдельных последовательностей оснований в двойной спирали ДНК? Насколько специфичны свойства этих последовательностей?

Одним из возможных путей создания промотора может быть использование определенной последовательности ДНК, являющейся сигналом для связывания РНК-полимеразы. Каждый промотор должен представлять собой (или, по крайней мере, включать в себя) такую последовательность. В бактериальном геноме минимальная длина участка, способного создать необходимый сигнал, составляет 12 пн (более короткие последовательности, вероятно, встречаются случайным образом довольно часто и могут создавать ложные сигналы). С увеличением размеров генома увеличивается и длина последовательности, обеспечивающей создание уникального сигнала распознавания. Последовательность промотора не обязательно должна быть непрерывной. Так, если консервативные последовательности разделяет определенное число любых пар оснований, совместная длина консервативных регионов может быть даже короче непрерывной промоторной последовательности (например, быть менее 12 пн), так как само разделяющее *расстояние* создает часть сигнала (даже если разделяющая *последовательность* не обладает никакой специфичностью).

Обнаружить характерные особенности ДНК, необходимые для связывания РНК-полимеразы, можно путем сравнения последовательности различных промоторов. Необходимая для связывания полимеразы нуклеотидная последовательность должна присутствовать во всех промоторах. Такую последовательность принято называть **консервативной.** Однако консервативная последовательность может слегка варьировать для разных промоторов. Каким же образом можно выявить достаточно консервативную и способную представлять собой уникальный сигнал последовательность ДНК?

Предполагаемые участки узнавания на ДНК можно представить в виде консенсусных последовательностей, составленных по принципу наиболее часто встречающихся в разных промоторах оснований. **Консенсусная последовательность** выводится на основе сравнения всех известных последовательностей некоторого региона. Для полученной таким способом консенсусной последовательности каждое конкретное основание должно наиболее часто встречаться в своем положении, причем большинство реальных последовательностей должно отличаться от консенсусной по небольшому числу позиций (1–2).

В результате анализа последовательностей промоторов *E. coli* была обнаружена удивительная особенность: среди 60 пн, взаимодействующих с РНК-полимеразой, *отсутствует сколько-нибудь протяженная консервативная последовательность.* Большая часть последовательности участка связывания может быть любой. Консервативными являются лишь отдельные короткие участки внутри промотора, они-то и могут быть сигналами узнавания для РНК-полимеразы. *Типичной чертой регуляторных сайтов (таких как промоторы) в прокариотических и эукариотических геномах является наличие отдельных коротких консенсусных последовательностей.*

Бактериальные промоторы характеризуются четырьмя (возможно, пятью) общими чертами: наличием стартовой точки, последовательности −10, последовательности −35, участка между положениями −10 и −35, и (иногда) наличием так называемого UP-элемента:

- Стартовой точкой, как правило (более чем в 90% случаев), служит пуриновое основание, обычно расположенное в середине последовательности CAT (однако консервативность этого триплета не настолько выражена, чтобы считать его обязательным сигналом).
- Выше стартовой точки находится область из 6 пн, обнаруживаемая почти во всех промоторах. Центр этой гексамерной последовательности обычно расположен на расстоянии примерно 10 пн выше стартовой точки транскрипции; для разных промоторов это расстояние варьирует от позиции −18 до позиции −9. По месту расположения эту гексамерную последовательность часто обозначают как **последовательность −10.** Последователь-

ность −10 имеет консенсус — *TATAAT*, и может быть представлена в виде

$T_{80}A_{95}T_{45}A_{60}A_{50}T_{96}$,

где цифрами показана частота (в процентах) встречаемости данного основания в данном положении (при описании значащих последовательностей положение, в котором с примерно равной вероятностью может находиться любое из оснований, обозначают как N). Если частота встречаемости основания указывает на его важное значение для взаимодействия с РНК-полимеразой, то мы вправе ожидать, что высококонсервативный димер TA в начале последовательности −10 и очень консервативный T в конце являются наиболее важными основаниями.

- Центр другой консервативной гексамерной последовательности расположен на расстоянии ~ 35 пн выше стартовой точки. Данный участок называют **последовательностью −35.** Эта последовательность — TTGACA, более детально выглядит как:

$T_{82}T_{84}G_{78}A_{65}C_{54}A_{45}$.

- Расстояние между положениями −35 и −10 варьирует от 16 до 18 пн у 90% промоторов; в отдельных случаях этот промежуток может уменьшаться до 15 пн или увеличиваться до 20 пн. *Хотя состав последовательности, находящейся между двумя консервативными областями, не важен, ее размер имеет критическое значение, поскольку определяет соответствие расположения консервативных участков структуре РНК-полимеразы.*
- Некоторые промоторы содержат A-T-богатую последовательность, расположенную еще выше регионов −10 и −35. Этот участок называют UP-элементом. Он служит для взаимодействия с α-субъединицей РНК-полимеразы. Как правило, эта регуляторная последовательность присутствует в промоторах генов с высоким уровнем экспрессии, таких как гены рРНК.

Оптимальный промотор состоит из −35 гексамерной последовательности, отделенной 17 пн от гексамерной последовательности −10, которая, в свою очередь, расположена на расстоянии 7 пн выше точки начала транскрипции. На РИС. 11.26 изображена структура такого промотора с указанием возможных отклонений от оптимума.

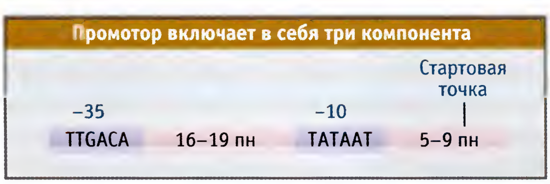


РИС. 11.26 Типичный промотор включает в себя три компонента: консенсусные последовательности −35 и −10 и стартовую точку

11.13 В результате мутаций эффективность работы промотора может быть повышена или понижена

Ключевые понятия:

- Ослабляющие мутации, уменьшая сходство промотора с консенсусными последовательностями, снижают его эффективность. Усиливающие мутации оказывают противоположный эффект.
- Мутации в последовательности -35, как правило, влияют на инициирующее связывание с РНК-полимеразой.
- Мутации в последовательности -10, как правило, влияют на процесс расплетания цепей ДНК, в ходе которого закрытый комплекс превращается в открытый.

Мутации являются основным источником информации о функционировании промоторов. Мутации в промоторе, влияя на уровень экспрессии гена (находящегося под контролем данного промотора), не изменяют кодируемого геном продукта. Большинство подобных мутаций было обнаружено в результате выделения бактериальных мутантов, у которых способность к транскрипции контролируемых промотором генов была либо совсем утрачена, либо находилась на очень низком уровне. Такие мутации называют **ослабляющими.** Реже возникают мутанты, обладающие повышенной транскрипцией с некоторого промотора. Это так называемые **усиливающие** мутации.

Важно помнить, что «ослабляющие» и «усиливающие» мутации характеризуют относительно *обычной* эффективности функционирования промотора. Однако этот параметр сам по себе широко варьирует для разных промоторов. Так, изменения, признанные ослабляющими для одного промотора, для другого могут быть недетектируемы (второй промотор даже в отсутствие мутаций может быть гораздо менее активен, чем мутантная форма первого). Информация, полученная в результате экспериментов *in vivo*, указывает лишь на природу обусловленного мутацией изменения (в целом).

Будет ли наиболее активным промотор, наиболее близкий к консенсусной структуре? Существует простое правило, которое гласит, что усиливающие мутации повышают сходство гексамерных последовательностей с консенсусными или приближают расстояние между ними к 17 пн. Ослабляющие мутации уменьшают сходство с консенсусными последовательностями или увеличивают разделяющее их расстояние. Ослабляющие мутации имеют тенденцию затрагивать наиболее консервативные позиции, что еще раз подтверждает важность этих положений в определении общей эффективности промотора. Существуют, однако, исключения из этого правила.

Чтобы оценить эффект мутации в промоторе, необходимо охарактеризовать взаимодействие РНК-полимеразы с промоторами мутантного и дикого типов *in vitro*. Существуют примерно 100-кратные вариации в эффективности связывания РНК-полимеразы с различными



РИС. 11.27 Последовательность −35 участвует в первичном узнавании промотора, а последовательность −10 в расплетании цепей ДНК, в результате чего закрытый комплекс превращается в открытый

промоторами *in vitro,* которые коррелируют с частотой инициации транскрипции с этих промотров *in vivo.* Можно экспериментально показать, на какое свойство промотора влияет мутация. Изменяет ли она сродство промотора к РНК-полимеразе? Может ли мутация, не нарушив способности фермента связываться с ДНК, мешать инициации? Изменяется ли под влиянием мутаций действие регуляторных факторов?

Определяя кинетические константы формирования закрытых комплексов и превращения их в открытые (как показано на рис. 11.19), можно выделить 2 отдельные стадии реакции инициации, на которые оказывают свой эффект мутации:

- Ослабляющие мутации в последовательности −35 уменьшают скорость образования закрытых комплексов (они понижают K_B), однако не влияют на превращение закрытых комплексов в открытые.
- Ослабляющие мутации в последовательности −10 не затрагивают процесса образования закрытых комплексов, однако замедляют их превращение в открытые (они понижают k_2).

На основе этих результатов была предложена модель, представленная на РИС. 11.27. Функция последовательности −35 заключается в обеспечении сигнала узнавания РНК-полимеразой, в то время как последовательность −10 отвечает за переход от закрытого комплекса к открытому. Мы можем рассматривать последовательность −35 в качестве «домена узнавания», а последовательность −10 — в качестве «домена расплетания».

Консенсусная последовательность в положении −10 состоит исключительно из оснований А и Т. Это позволяет предположить, что ее функционирование связано с плавлением ДНК и образованием одноцепочечных участков. Для разрыва водородных связей в парах А-Т требуется меньше энергии, чем для разрыва связей в парах G-C. Из этого следует, что для расхождения цепей в участке, состоящем из пар А-Т, требуется меньшее количество энергии.

Последовательность, непосредственно прилегающая к стартовой точке, также влияет на инициацию. Транскрибируемый участок, находящийся сразу за стартовой точкой по ходу транскрипции (от положения +1 до положения +30), влияет на скорость, с которой РНК-полимераза освобождает промотор (для следующего цикла инициации), внося тем самым вклад в общую эффективность работы промотора. Таким образом, эффективность работы промотора определяется не только консенсусными последовательностями −35 и −10.

«Типичный» промотор использует последовательности −35 и −10 для взаимодействия с РНК-полимеразой, однако у ряда промоторов одна из рассмотренных последовательностей может отсутствовать. По крайней мере в некоторых из этих случаев промотор сам по себе не узнается РНК-полимеразой; реакция распознавания требует участия дополнительных белков, с помощью которых восполняется отсутствие консенсусных последовательностей.

11.14 РНК-полимераза связывает молекулу ДНК с одной стороны

Ключевые понятия:

- Консенсусные последовательности −35 и −10 являются точками контакта РНК-полимеразы и промотора.
- Точки контакта расположены на молекуле ДНК с одной стороны

Способность РНК-полимеразы (или любого другого белка) взаимодействовать с определенными участками ДНК можно исследовать с помощью метода **футпринтинга**. Для этого связанную белком ДНК *частично* расщепляют с помощью эндонуклеаз, неспецифично гидролизующих любые фосфодиэфирные связи нуклеиновой кислоты. Условия реакции подбирают так, чтобы ДНК разрушалась не до отдельных нуклеотидов, а на фрагменты, причем благодаря случайному распределению сайтов расщепления в смеси окажутся молекулы любой длины. Местоположение подвергшихся гидролизу связей определяют благодаря ранее (до расщепления) введенной в одну из цепей ДНК и только с одного конца метке (этот же принцип лежит в основе одного из методов секвенирования ДНК: частичное расщепление в определенных сайтах молекул ДНК с введенной в них концевой меткой с формированием фрагментов определенной длины).

Как показано на РИС. 11.28, в результате электрофореза обработанных нуклеазой молекул ДНК фрагменты

будут разделяться в геле соответственно своей длине. Каждому фрагменту (несущему радиоактивную концевую метку) будет соответствовать отдельное пятно (полоса) на радиоафтографе. Положение полосы в геле будет зависеть от длины фрагмента (от меченого края до места расщепления эндонуклеазой). Короткие фрагменты движутся в геле быстрее, поэтому фрагменты с меньшим расстоянием от меченого края до места разрыва окажутся внизу геля.

В не связанной с белком ДНК *любая* фосфодиэфирная связь может быть гидролизована. Однако в случае взаимодействия с белком участок ДНК может быть экранирован связавшимся белком и защищен от действия нуклеазы (во всех молекулах). В эксперименте параллельно проводят две реакции: с контрольной (свободной ДНК) и с ДНК, провзаимодействовавшей с белком. *В участке ДНК, защищенном от действия нуклеаз белком, разрыва связей не происходит.*

В контрольной же ДНК доступны все связи. В результате этого образуется серия полос, отличающихся по длине на одно основание. На рисунке представлена 31 полоса. В экранированном фрагменте часть связей не может быть гидролизована, поэтому полосы определенного размера (соответствующие экранированному региону) не образуются. Отсутствующие на рисунке полосы от 9 до 18 соответствуют белок-связывающему сайту на ДНК (расположенному, соответственно, от 9-го до 18-го основания, считая от меченого конца молекулы ДНК). Сравнивая электрофоретические профили, полученные при гидролизе контрольной и опытной ДНК, с результатами реакции секвенирования (проводимой параллельно), становится возможным «прочитать» нуклеотидную последовательность непосредственно в белок-связывающем участке.

Как уже было сказано (см. рис. 11.20), РНК-полимераза связывает регион, расположенный между положениями –50 и +20. Точки контакта РНК-полимеразы и промотора могут быть определены с помощью одного из вариантов метода футпринтинга, в основе которого лежит обработка комплекса РНК-полимераза-промотор реагентами, модифицирующими определенные основания в ДНК. Схемы проведения такого эксперимента могут быть следующими:

- ДНК может быть модифицирована до взаимодействия с РНК-полимеразой. Если модификация повлияет на способность фермента связываться с ДНК, это будет означать, что модификация затронула основания, играющие важную роль в этом взаимодействии.
- В случае когда обрабатывают комплекс РНК-полимераза–ДНК, после модификации проводят сравнение результатов гидролитического расщепления модифицированного комплекса, свободной ДНК и немодифицированного комплекса. Исчезнувшие полосы будут соответствовать фрагментам, которые фермент экранирует от модификации. Полосы, интенсивность свечения которых увеличится, будут соответствовать фрагментам молекулы ДНК, которые в результате изменения конформации комплекса становятся более доступны действию нуклеаз.

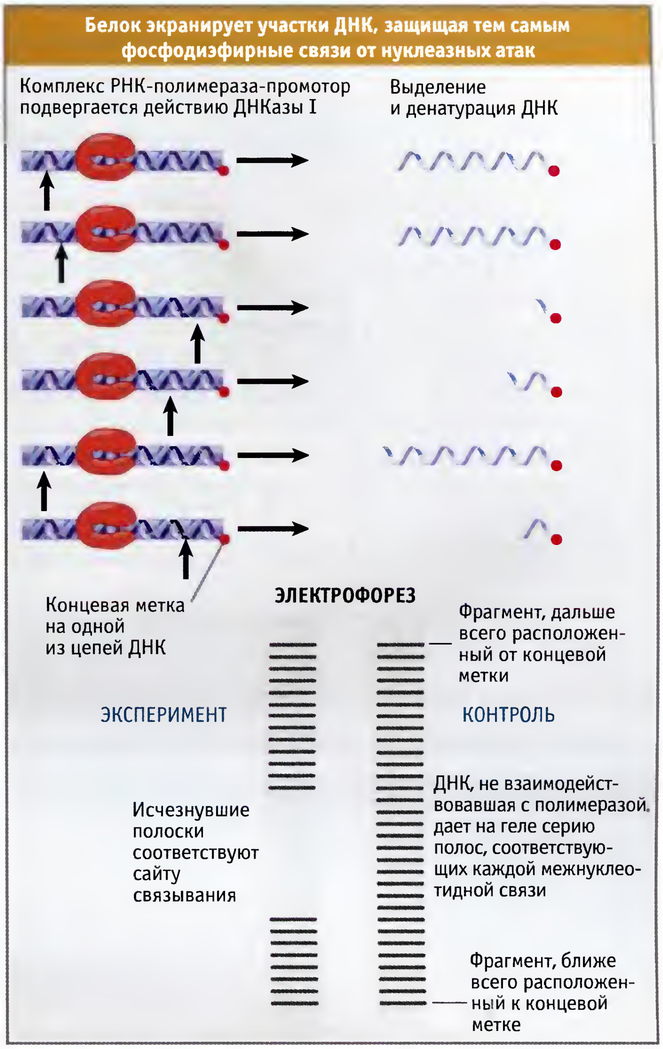


РИС. 11.28 С помощью метода отпечатков (футпринтинга) на ДНК были определены белок-связывающие сайты

Изменение чувствительности различных участков ДНК к нуклеазной атаке отражает геометрию комплекса ДНК-фермент. Схема строения данного комплекса для типичного промотора приведена на РИС. 11.29. Большинство точек, по которым происходит взаимодействие РНК-полимеразы с ДНК, находятся в областях –35 и –10. В пределах этих областей участвующие в связывании группы оснований либо предотвращают связывание белка (если они предварительно были модифицированы), либо проявляют повышенную (или пониженную) чувствительность к модификации в результате связывания фермента. Точки контакта белка и ДНК могут частично не совпадать с сайтами, затронутыми мутациями; однако значимые мутации присутствуют в одних и тех же ограниченных областях.

Следует отметить, что точками контакта в различных промоторах являются тождественные (по положению) *позиции*, даже если они содержат разные основания. Из этого можно сделать вывод о существовании общего механизма связывания РНК-полимеразы, причем ре-

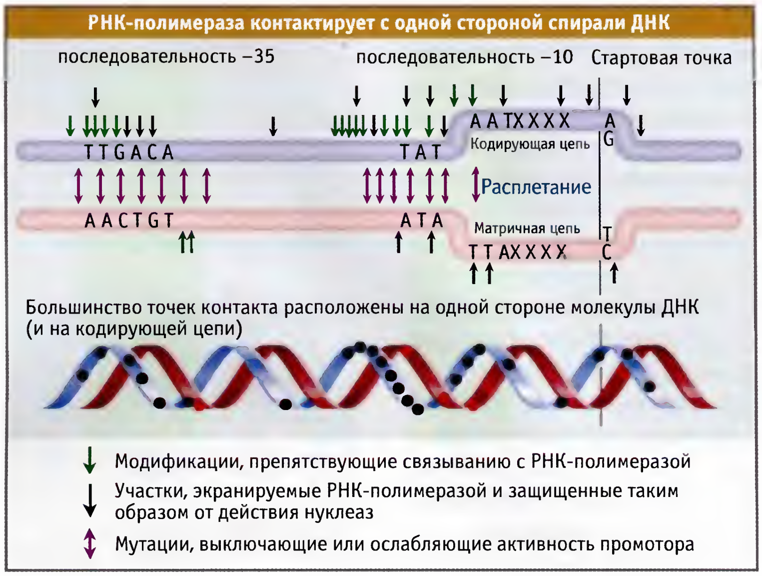


РИС. 11.29 Точки контакта РНК-полимеразы с промотором расположены на одной стороне спирали ДНК

акция не зависит от наличия конкретных оснований в какой-либо из точек контакта. Эта модель объясняет, почему некоторые точки контакта не являются сайтами, чувствительными к мутациям. Кроме того, не каждая значащая мутация затрагивает точку контакта ДНК и фермента; мутации могут оказывать влияние на соседние основания, не соприкасающиеся с ферментом.

Два описанных выше экспериментальных подхода дают возможность исследовать разные типы контактов между ДНК и ферментом. Эксперименты с предварительной модификацией ДНК выявляют *только* сайты, защищенные ферментом от последующих модификаций (но не обязательно находящиеся в непосредственном контакте), тогда как в случае эксперимента с модификацией определяются сайты, которые действительно образуют контакты в двойном комплексе. Так как экранированные ферментом участки включают в себя все позиции, участвующие в узнавании, и еще ряд дополнительных близлежащих позиций, то, по-видимому, фермент сначала узнает некий набор сайтов, необходимых для его «посадки», и только после этого образует более обширные контакты.

Расплетенная область ДНК, входящая в состав двойного комплекса, может быть обнаружена с помощью химических реагентов, выявляющих одноцепочечную ДНК. При разделении цепей ДНК, неспаренные основания становятся чувствительны к реагентам, от действия которых они защищены в двухцепочечном состоянии. В результате таких экспериментов было выяснено, что регион первичного расплетания цепей ДНК находится между позициями −9 и +3. Область, расплетаемая в реакции инициации транскрипции, включает участок, начинающийся от правого конца последовательности −10 и заканчивающийся сразу за стартовой точкой.

При пространственном рассмотрении все точки контакта, находящиеся выше последовательности −10, расположены на одной и той же поверхности ДНК. Это проиллюстрировано на схеме, приведенной в нижней части рис. 11.29, где большинство точек контакта расположены на одной стороне двойной спирали ДНК. Причем большая часть из них лежит на кодирующей цепи. Эти основания, вероятно, узнаются при первоначальном образовании закрытого двойного комплекса, что обеспечивает реакцию узнавания между РНК-полимеразой и ДНК путем образования контакта между ферментом и одной из поверхностей ДНК. После того как расплетание ДНК началось, становится возможным узнавание и связывание со следующими сайтами, расположенными на другой стороне ДНК.

11.15 Сверхспирализация ДНК влияет на процесс транскрипции

Ключевые понятия:

- Отрицательная сверхспирализация повышает эффективность работы некоторых промоторов, способствуя расплетанию цепей ДНК.
- В ходе транскрипции перед ферментом возникает положительная сверхспирализация ДНК, а позади фермента — отрицательная сверхспирализация. Оба варианта сверхспирализации должны быть релаксированы с помощью гираз и топоизомераз.

О важном значении расплетания цепей при инициации транскрипции свидетельствует эффект, оказываемый сверхспирализацией. Как прокариотические, так и эукариотические РНК-полимеразы *in vitro* более эффективно инициируют транскрипцию на сверхспирализованных матрицах. Это объясняется тем, что для первоначального плавления ДНК в инициирующем

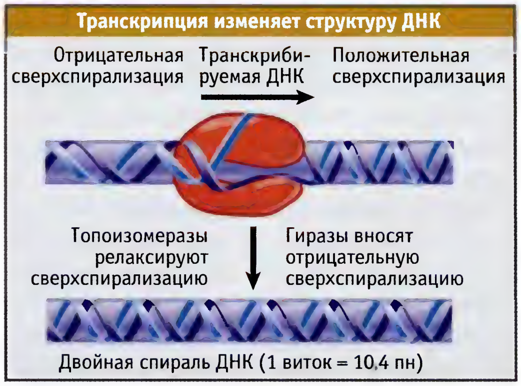


РИС. 11.30 **По мере движения РНК-полимеразы впереди фермента формируется участок ДНК, скрученный более сильно, а позади — участок ДНК с частично расплетенными цепями**

комплексе сверхспирализованная структура требует гораздо меньших затрат энергии.

Степень сверхспирализации влияет на эффективность работы некоторых промоторов (отрицательная сверхспирализация облегчает транскрипцию, что является характерной чертой процесса транскрипции). Природа эффекта, оказываемого сверхспирализацией на реакцию инициации, известна. Однако не все промоторы чувствительны к сверхспирализации. Одно из объяснений этого факта состоит в том, что зависимость каждого промотора от степени сверхспирализации определяется его нуклеотидной последовательностью. Некоторые промоторы имеют последовательности, которые плавятся легче (и поэтому в меньшей степени зависят от сверхспирализации), тогда как другие обладают тугоплавкими последовательностями (и в большей степени зависят от сверхспирализации). Другим объяснением могут быть различия в степени сверхспирализации для разных регионов бактериальной хромосомы, в результате чего важным становится местоположение промотора в геноме.

Сверхспирализация, безусловно, оказывает важный эффект на процессы транскрипции. Как показано на рис. 11.4, транскрибирование молекулы ДНК сопровождается расплетанием и последующим восстановлением двойной спирали. Следовательно, транскрипционный комплекс должен иметь возможность вращаться вокруг молекулы ДНК либо сама ДНК должна вращаться вокруг своей оси. На РИС. 11.30 изображены последствия считывания некого региона ДНК согласно модели транскрипции, получившей название модели *двойного домена*. Двигаясь вдоль двойной спирали, РНК-полимераза формирует положительную сверхспирализацию впереди себя (участок ДНК как бы закручивается еще сильнее) и создает отрицательную сверхспирализацию за собой (формируя участок ДНК с частично расплетенными цепями). Для каждого витка спирали, развернутого РНК-полимеразой, впереди формируется 1 положительный сверхвиток, а позади формируется 1 отрицательный сверхвиток.

Транскрипция оказывает значительный эффект на локальную структуру молекулы ДНК. В связи с этим фермент гираза (способный создавать отрицательную сверхспирализацию) и фермент топоизомераза I (способный релаксировать отрицательную сверхспирализацию) необходимы для корректировки структуры ДНК перед полимеразой и позади нее соответственно. Ингибирование активностей гиразы и топоизомеразы I нарушает процессы сверхспирализации ДНК. Так, например, в дрожжах, дефицитных по ферменту, релаксирующему отрицательные суперспирали, плотность участков отрицательной сверхспирализации в транскрибируемом регионе удваивается. На основе этих данных можно предположить, что транскрипция ответственна за формирование значительной части сверхспирализованных участков ДНК в клетке.

Похожая ситуация наблюдается и при репликации, где расплетание цепей ДНК и движение репликативной вилки происходят таким образом, чтобы отдельные цепи ДНК могли служить матрицей для синтеза дочерних цепей. (Пути преодоления топологических трудностей, возникающих в ходе данной реакции, приведены на рис. 19.20.)

11.16 Инициация на данном промоторе зависит от типа сигма-фактора

Ключевые понятия:

- *E.coli* имеет несколько сигма-факторов, каждый из которых обеспечивает связывание РНК-полимеразы с отдельной группой промоторов (со схожими -35 и -10 последовательностями).
- σ^{70} является основным фактором, инициирующим транскрипцию, другие сигма-факторы активируются только при определенных условиях.

Разделение функций между минимальным ферментом (осуществляющим элонгацию нуклеотидной цепи) и сигма-фактором (ответственным за распознавание сайта связывания), поднимает вопрос о существовании нескольких типов сигма-факторов, специфичных к различным классам промоторов. На РИС. 11.31 изобра-

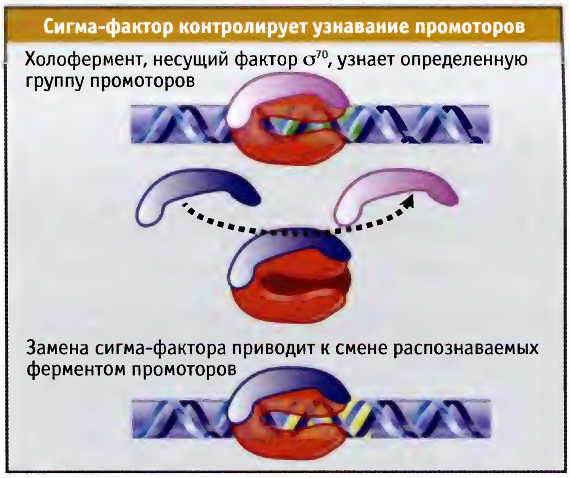


РИС. 11.31 **Сигма-фактор взаимодействует с минимальным ферментом, определяя набор промоторов, с которых будет происходить инициация транскрипции**

E. coli обладает несколькими сигма-факторами

Ген	Фактор	Условия использования
rpoD	σ^{70}	основной
rpoS	σ^{S}	стресс
rpoH	σ^{32}	тепловой шок
rpoE	σ^{E}	тепловой шок
rpoN	σ^{54}	азотное голодание
fliA	σ^{28} (σ^{F})	синтез флагеллярных структур

РИС. 11.32 Кроме фактора σ^{70}, *E. coli* обладает несколькими альтернативными сигма-факторами, которые активируются в ответ на определенные условия окружающей среды (цифра в названии фактора показывает его массу)

жен принцип функционирования системы, в которой замена сигма-фактора влияет на выбор промотора.

На РИС. 11.32 перечислены альтернативные сигма-факторы, обнаруженные в клетках *E. coli* и используемые бактерией при изменении условий окружающей среды (названия даны в соответствии с молекулярным весом белка или согласно названию гена). Основным фактором, отвечающим за транскрипцию большинства генов в нормальных условиях, является σ^{70}. Альтернативные сигма-факторы σ^{S}, σ^{32}, σ^{E} и σ^{54} активируются в ответ на изменение внешних условий; σ^{28} задействован в экспрессии флагеллярных генов в нормальных условиях, однако уровень его экспрессии зависит от изменения условий окружающей среды. Все сигма-факторы (за исключением σ^{54}) принадлежат к одному белковому семейству и функционируют по сходному механизму.

Изменения температуры являются наиболее частым случаем внешних воздействий. Многие организмы, как про-, так и эукариоты, реагируют на такие изменения схожим образом. При повышении температуры синтез обычного набора белков понижается или прекращается и начинается синтез белков, являющихся продуктами **генов теплового шока.** Белки теплового шока участвуют в защите клетки от стресса (их синтез запускается не только в ответ на повышение температуры). Некоторые из этих белков являются шаперонами. У *E. coli* уровень экспрессии семнадцати генов белков теплового шока опосредован изменениями в процессе транскрипции соответствующих генов. В переключении аппарата белкового синтеза на продукцию белков теплового шока участвует регуляторный ген *rpoH*. Продукт этого гена — белок σ^{32}, функционирует как альтернативный фактор, инициируя транскрипцию генов белков теплового шока.

При повышении температуры концентрация σ^{32} возрастает, запуская синтез белков теплового шока, а при понижении температуры активность σ^{32} падает. Основным сигналом, индуцирующим продукцию σ^{32}, является накопление в клетке несвернутых (частично денатурированных в результате действия высоких температур) белков. Сам белок σ^{32} короткоживущий, что определяет возможность не только его быстрого накопления, но и столь же быстрого разрушения. Возможна конкуренция между белками σ^{70} и σ^{32} за доступный для связывания базовый фермент, поэтому набор генов, транскрибируемых во время теплового шока, зависит от баланса между этими двумя белками. Смена сигма-факторов является широко распространенным способом контроля экспрессии генов у бактерий, а продукция новых сигма-факторов может быть основой многих регуляторных механизмов.

Синтез сигма-фактора σ^{S} индуцируется, когда фаза роста бактериальной культуры переходит в стационарную фазу (а также в ряде других стрессовых условий). Контроль продукции фактора осуществляется на двух уровнях: во-первых, трансляция мРНК гена *rpoS* повышается при низких температурах или высокой осмолярности. Во-вторых, протеолиз белка ингибируется углеродным голоданием (типичная черта стационарной фазы роста культуры) и высокими температурами.

Другая группа регуляторных генов, участвующих в ответе клетки на температурные воздействия, контролируется фактором σ^{E}. Этот фактор синтезируется в ответ на еще более резкие температурные сдвиги (чем фактор σ^{32}), его синтез индуцируется в результате накопления денатурированных белков в периплазматическом пространстве или на внешней мембране. Продукция σ^{E} контролируется механизмами, представленными на РИС. 11.33. Фактор σ^{E} связан с белком (RseA), локализованным во внутренней мембране клеточной оболочки, в результате чего он не может активировать транскрипцию. Накопление денатурированных белков активирует протеазу (DegS) периплазматического пространства,

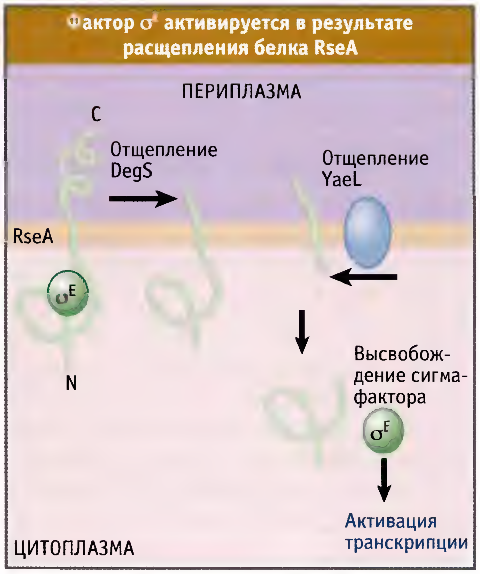


РИС. 11.33 RseA — белок внутренней мембраны. Его цитоплазматический домен связывает фактор σ^{E}. Протеазы периплазматического пространства, а затем цитоплазмы, последовательно отщепляют от RseA концевые фрагменты. Отщепление N-концевого участка ведет к высвобождению σ^{E}

отщепляющую С-концевой участок белка RseA. В результате такого отщепления активируется другой белок внутренней мембраны (YaeL), который, в свою очередь, отщепляет N-концевой участок RseA. Это ведет к высвобождению фактора σ^E, запускающего транскрипцию. Конечным результатом накопления денатурированных белков на периферии бактериальной клетки является активизация работы генов, контролирующих продукцию сигма-факторов.

Можно провести две интересные параллели между данным регуляторным механизмом и работой других регуляторных циклов. Ответ эукариотической клетки на накопление денатурированных белков опосредован механизмом, согласно которому денатурированный белок (внутри эндоплазматического ретикулума) активирует мембранный белок — эндонуклеазу, расщепляющую РНК, что в конечном итоге приводит к изменению сплайсинга и продукции транскрипционных факторов (см. разд. 26.17 «Каскад реакций, запускаемый в ответ на накопление белков, связан с процессом сплайсинга тРНК»). Более тесную параллель можно провести с первыми обнаруженными случаями, когда расщепление мембранного белка активировало непосредственно транскрипционный фактор. Транскрипционный фактор синтезируется в виде мембранного белка, а уровень стеролов в мембране контролирует процесс активации протеаз, отщепляющих транскрипционный фактор от цитозольного домена белка.

В условиях азотного голодания клеткой используется другой сигма-фактор. Клетки *E.coli* содержат небольшое количество фактора σ^{54}, который активируется при отсутствии ионов аммония в культуральной среде. В этих условиях активность генов переключается таким образом, чтобы сделать возможным использование альтернативных источников азота. Аналоги этого сигма-фактора обнаружены у широкого спектра бактерий, что указывает на закрепленный в ходе эволюции механизм ответа на стресс.

Другим примером эволюционно консервативного сигма-фактора является фактор σ^F, присутствующий в клетке в небольших количествах и инициирующий транскрипцию генов, вовлеченных в хемотаксис и формирование флагеллярных структур. Его аналогом у *Bacillus subtilis* является фактор σ^D, отвечающий за транскрипцию флагеллярных генов и генов органов движения; факторы, характеризующиеся специфичностью к этому же набору промоторов, присутствуют у многих видов бактерий.

Таким образом, каждый сигма-фактор отвечает за связь РНК-полимеразы с определенным набором промоторов. Анализируя последовательности этих промоторов, можно заметить, что каждая группа характеризуется наличием уникальных структурных элементов. Действительно, промоторы каждого типа узнаются только комплексом РНК-полимеразы с подходящим сигма-фактором. Благодаря идентификации генов, кодирующих сигма-факторы *E.coli,* а также изучению сигма-факторов, вовлеченных в споруляцию у *Bacillus subtilis* (см. разд. 11.19 «Спорообразование контролируется сигма-факторами»), можно сформулировать общие правила, по которым происходит распознавание промотора.

Сигма-факторы распознают промоторы по консенсусным последовательностям

Ген	Фактор	Последовательность −35	Расстояние между консенсусными последовательностями	Последовательность −10
rpoD	σ^{70}	TTGACA	16–18 пн	TATAAT
rpoH	σ^{32}	CCCTTGAA	13–15 пн	CCCGATNT
rpoN	σ^{54}	CTGGNA	6 пн	TTGCA
fliA	σ^{28} (σ^F)	CTAAA	15 пн	GCCGATAA
sigH	σ^H	AGGANPuPu	11–12 пн	GCTGAATCA

РИС. 11.34 Разные сигма-факторы *E. coli* узнают промоторы с разными консенсусными последовательностями

Отличительным свойством промоторов, обеспечивающих их узнавание конкретным холоферментом, является их *одинаковый размер и расположение относительно стартовой точки, при этом консервативные последовательности всегда расположены в районе позиций −35 и −10* (фактор σ^{54} является исключением, консенсусные последовательности промоторов, которые он распознает, расположены ближе друг к другу в положениях −24 и −12; см. разд. 11.17 «Сигма-факторы взаимодействуют непосредственно с ДНК»). Как показано на РИС. 11.34, консенсусные последовательности разных групп промоторов могут отличаться друг от друга последовательностью −35, последовательностью −10, а также обеими этими позициями. Это означает, что фермент, содержащий в своем составе определенный сигма-фактор, может распознавать только свой набор промоторов, при этом инициация транскрипции с разных групп промоторов взаимно исключается. Замена одного сигма-фактора на другой выключает транскрипцию ранее экспрессируемых генов и включает транскрипцию новых. (Гены, имеющие несколько промоторов с разным набором консенсусных последовательностей, могут транскрибироваться РНК-полимеразой, содержащей в своем составе разные сигма-факторы.)

11.17 Сигма-факторы взаимодействуют непосредственно с ДНК

Ключевые понятия:

- Изменение конформации σ^{70}, ассоциированного с минимальным ферментом, приводит к высвобождению ДНК-связывающих регионов фактора.
- σ^{70} связывает как −35, так и −10 последовательности.

Идентификация серии консенсусных последовательностей −35 и −10, распознаваемых холоферментом, содержащим в своем составе разные сигма-факторы, позволила

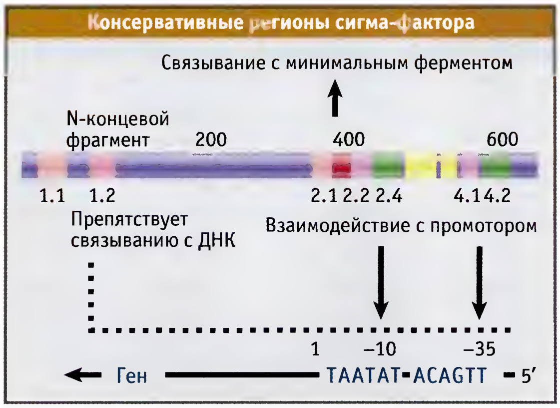


РИС. 11.35 На генетической карте фактора σ^{70} *E. coli* показаны консервативные регионы. Регионы 2.1 и 2.2 контактируют с минимальным ферментом, 2.3 участвует в расплетании ДНК, 2.4 и 4.2 образуют контакты с последовательностями промотора –35 и –10. N-концевой участок не дает регионам 2.4 и 4.2 связываться с ДНК в отсутствие минимального фермента

сделать вывод, что субъединица сигма-фактора непосредственно взаимодействует с ДНК в этих регионах. В результате был сформулирован общий принцип взаимодействия сигма-фактора и минимального фермента, согласно которому сигма-фактор ориентирован таким образом, чтобы образовывать тесные контакты с промоторной последовательностью вблизи положений –35 и –10.

Ярким свидетельством прямого контакта между сигма-фактором и промотором в области последовательностей –35 и –10 являются мутации сигма-фактора, способные супрессировать мутации в консенсусных последовательностях. В случае когда мутации в определенных положениях промоторной последовательности препятствуют ее узнаванию РНК-полимеразой, компенсаторные мутации в сигма-факторе позволяют РНК-полимеразе взаимодействовать с мутантным промотором. Наиболее вероятное объяснение этого факта заключается в том, что подходящие пары оснований в ДНК контактируют с соответствующими аминокислотными заменами в сигма-факторе.

Сравнение последовательностей нескольких бактериальных сигма-факторов позволило выявить консервативные регионы. Их местоположение в факторе σ^{70} *E. coli* представлено на РИС. 11.35. Из структуры псевдокристалла фрагмента сигма-фактора, выделенного из бактерии *Thermus aquaticus*, видно, что эти участки формируют в белке три независимых домена: домен σ_2 включает в себя регионы 1.2–2.4, домен σ_3 — 3.0–3.1, домен σ_4 — 4.1–4.2.

Как показано на рис. 11.35, два коротких сегмента регионов 2 и 4 (названные 2.4 и 4.2) участвуют во взаимодействии с промотором в районах –10 и –35 соответственно. Оба этих региона белка формируют небольшие α-спиральные участки. Эксперименты с гетеродуплексами показали, что σ^{70} образует контакты с основаниями исключительно кодирующей цепи и сохраняет эту связь после того, как в области контакта произойдет расплетание цепей ДНК. Основываясь на этих данных, можно предположить, что сигма-фактор играет важную роль в реакции плавления двухцепочечной структуры ДНК.

Использование белком α-спиральных участков для распознавания двухцепочечной ДНК является общим свойством всех сигма-факторов (см. разд. 14.11 «Репрессоры связываются с ДНК посредством мотива спираль-поворот-спираль»). Аминокислоты, удаленные друг от друга на три-четыре положения, располагаются на одной поверхности α-спирали, что способствует их взаимодействию с соседними основаниями цепи ДНК. Как показано на РИС. 11.36, аминокислоты, расположенные на одной поверхности α-спирали региона 2.4, взаимодействуют с основаниями в позициях от –12 до –10 промоторной последовательности –10.

Регион 2.3 похож на белки, связывающие одноцепочечные нуклеиновые кислоты, и участвует в реакции расплетания ДНК. Регионы 2.1 и 2.2 (являющиеся наиболее консервативными участками сигма-фактора) вовлечены во взаимодействие с минимальным ферментом. Это позволяет предположить, что все сигма-факторы связывают минимальный фермент в одних и тех же участках, что обеспечивает конкуренцию между сигма-факторами за связь с полимеразой.

N-концевой участок фактора σ^{70} выполняет важные регуляторные функции. При его удалении укороченный белок приобретает способность специфично связываться с промоторами (без минимального фермента). Это позволяет предположить, что N-концевой регион функционирует в качестве автоингибирующего домена, блокирующего ДНК-связывающие участки свободного σ^{70}. Связываясь с минимальным ферментом, сигма-фактор меняет свою конформацию таким образом, чтобы снять ингибирующий эффект N-концевого участка и сделать возможным взаимодействие контактных доменов с молекулой ДНК.

На РИС. 11.37 схематично показаны конформационные изменения сигма-фактора при формировании открытого комплекса. При ассоциации сигма-фактора

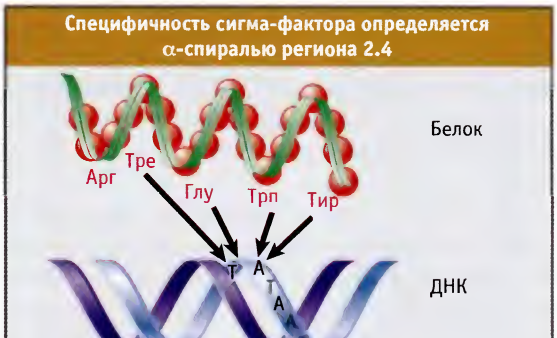


РИС. 11.36 Аминокислоты α-спирали региона 2.4 фактора σ^{70} контактируют со специфическими основаниями кодирующей цепи промоторной последовательности –10

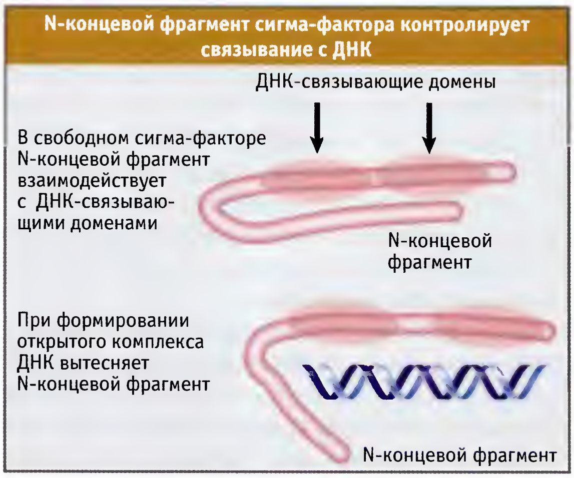


РИС. 11.37 N-концевой фрагмент сигма-фактора блокирует ДНК-связывающий участок от связывания с ДНК. При формировании открытого комплекса, N-конец отодвигается от ДНК-связывающих доменов на 20 Å, сами ДНК-связывающие домены при этом удаляются друг от друга на 15 Å



РИС. 11.38 Сигма-фактор имеет протяженную структуру, которая растягивается вдоль поверхности минимального фермента при формировании холофермента

с минимальным ферментом N-концевой домен отходит от ДНК-связывающих доменов примерно на 20 Å. Сами ДНК-связывающие домены при этом удаляются друг от друга примерно на 15 Å, приобретая, по-видимому, конформацию, наиболее удобную для формирования контактов с ДНК. Мутации в последовательностях −10 и −35 препятствуют связыванию фактора σ^{70} (лишенного N-концевого участка) с ДНК. Это позволяет предположить, что σ^{70} контактирует одновременно с обеими последовательностями. Для того чтобы это стало возможным, сигма-фактор должен иметь удлиненную структуру, протягивающуюся примерно на 68 Å — расстояние, равное длине двух витков спирали ДНК.

В свободном холоферменте N-концевой домен расположен в активном центре фермента и экранирует собой место, которое впоследствии при формировании транскрипционного комплекса займет ДНК. Когда холофермент формирует с ДНК открытый комплекс, N-концевой домен сигма-фактора вытесняется из активного центра. Взаимодействия этого участка молекулы с остальными регионами фермента очень гибкие; характер этих взаимодействий меняется и когда сигма-фактор связывается с минимальным ферментом, и когда холофермент связывается с ДНК.

Сравнение структур псевдокристаллов минимального фермента и холофермента показало, что сигма-фактор располагается главным образом на поверхности минимального фермента. Как показано на РИС. 11.38, он имеет протяженную структуру, распространяющуюся за пределы ДНК-связывающего сайта. Это позволяет сигма-фактору контактировать с ДНК во время первичного связывания. Для того чтобы войти в активный центр фермента, спираль ДНК сдвигается на 16 Å от точки первичного связывания. РИС. 11.39 иллюстрирует этот сдвиг на поперечном разрезе молекулы ДНК (вдоль оси спирали).

Интересная особенность была обнаружена в функционировании фактора σ^{54}. Он способствует связыванию РНК-полимеразы с промоторами, содержащими определенную консенсусную последовательность с одним консервативным элементом, расположенным в позиции −12, и другим, расположенным в районе −24 (приведен в колонке «−35» рис. 11.34). Таким образом, геометрия комплекса полимераза-промотор отличается в зависимости от сигма-фактора, входящего в состав фермента. Еще одна особенность этого сигма-фактора состоит в том, что для эффективной транскрипции, опосредуемой фактором σ^{54}, требуются



РИС. 11.39 Сначала ДНК контактирует с сигма-фактором (розовый) и минимальным ферментом (серый). Затем она продвигается глубже в структуру минимального фермента для того, чтобы последовательность −10 могла образовать контакты с белком. После высвобождения сигма-фактора ширина канала, в котором расположена ДНК, увеличивается. Воспроизведено с материалов Vassylyev, D.G., et. al. *Nature*. 2002. 417: 712–719. Фотография любезно предоставлена Shigeyuki Yokoyama, The University of Tokio

дополнительные активаторы, взаимодействующие с участками, расположенными довольно далеко от промотора. Регуляторные сайты всех других типов бактериальных промоторов расположены в непосредственной близости от самого промотора. Механизм функционирования фактора σ^{54} отличается от работы остальных сигма-факторов. Наиболее характерной особенностью является его способность связывать ДНК независимо от минимального фермента. В этом отношении σ^{54} больше похож на эукариотические транскрипционные факторы (обсуждаемые в гл. 24 «Промоторы и энхансеры»), чем на типичные прокариотические регуляторы (обсуждаемые в гл. 12 «Оперон»).

11.18 Сигма-факторы организованы в каскады

Ключевые понятия:

- Каскад сигма-факторов формируется в том случае, если один сигма-фактор необходим для транскрипции гена, кодирующего следующий сигма-фактор.
- Ранние гены фага SP01 транскрибируются РНК-полимеразой клетки-хозяина.
- Один из ранних генов кодирует сигма-фактор, заставляющий РНК-полимеразу транскрибировать средние гены фага.
- Два средних гена кодируют субъединицы сигма-фактора, обеспечивающего транскрипцию РНК-полимеразой поздних генов.

В клетке бактерии *B. subtilis* сигма-факторы широко используются для контроля инициации транскрипции, для этого микроорганизма известно порядка 10 различных σ-факторов. Некоторые из них присутствуют в обычном состоянии; другие появляются только в определенных условиях — при фаговой инфекции или при переходе от стадии вегетативного роста клетки к стадии споруляции.

Основная РНК-полимераза, найденная у *B. subtilis* и вовлеченная в процессы нормального вегетативного роста клетки, имеет структуру, схожую со структурой полимеразы *E. coli* $\alpha_2\beta\beta'\sigma$. Этот сигма-фактор (называемый σ^{43} или σ^{A}) распознает промоторы с такими же консервативными последовательностями, как и фактор σ^{70} *E. coli*. Несколько вариантов РНК-полимеразы, содержащей в своем составе другие сигма-факторы, были обнаружены в гораздо меньших количествах. Разные варианты фермента распознают разные промоторы в районе консенсусных последовательностей –35 и –10.

Для фаговой инфекции характерно выключение экспрессии обычной группы генов и включение новых генов. Практически во всех случаях при развитии фага в ходе инфекционного цикла происходит переключение транскрипции. Это переключение приводит либо к синтезу фаговой РНК-полимеразы, либо к контролю бактериальной РНК-полимеразы, осуществляемому вспомогательными факторами фага. Хорошо охарактеризованным примером контроля транскрипции посредством продукции нового сигма-фактора является инфицирование бактерии *B.subtilis* фагом SPO1.

В зависимости от того, какие гены экспрессируются, инфекционный цикл фага SPO1 можно разделить на три стадии. Сразу же после инфицирования бактериальной клетки транскрибируются так называемые **ранние гены** фага. Через 4–5 минут транскрипция ранних генов прекращается и начинается транскрипция **средних генов**. На 8–12 мин транскрипция средних генов заменяется транскрипцией **поздних генов**.

Ранние гены фага транскрибируются холоферментом бактерии-хозяина. Они, по существу, не отличаются от хозяйских генов, чьи промоторы способна узнавать РНК-полимераза состава $\alpha_2\beta\beta'\sigma^{43}$. Экспрессия ранних генов фага необходима для транскрипции средних и поздних генов. Три регуляторных гена, обозначаемых *28*, *33* и *34*, контролируют очередность транскрипции. Функции этих генов представлены на РИС. 11.40. Регуляция осуществляется **каскадно**, при этом бактериальный фермент сначала транскрибирует ранний ген, чей продукт необходим для экспрессии средних генов. В свою очередь, продукты двух средних генов необходимы для транскрипции поздних генов.

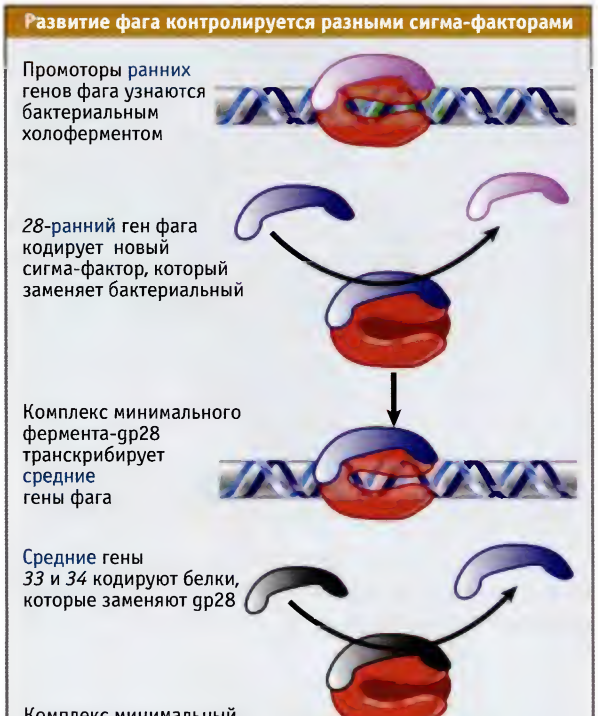


РИС. 11.40 Транскрипция генов фага SP01 на разных стадиях инфекционного цикла контролируется путем смены сигма-факторов, что приводит к изменению специфичности инициации

Мутанты, дефектные по раннему гену *28*, не способны транскрибировать средние гены. Продукт этого гена (обозначаемый gp28) представляет собой белок с молекулярной массой 26 кДа, заменяющий в минимальном ферменте сигма-фактор хозяина. *Эта замена является единственным событием, необходимым для переключения транскрипции с ранних генов на средние.* Образующийся при этом холофермент не может продолжать транскрибировать гены клетки-хозяина, однако приобретает способность специфически транскрибировать средние гены фага. Каким образом gp28 заменяет σ^{43} и какова при этом судьба бактериального сигма-фактора, неизвестно.

Два средних гена принимают участие в следующем переключении транскрипции. Мутация в гене *33* или *34* предотвращает транскрипцию поздних генов. Продукты этих генов формируют димер, который замещает gp28 в холоферменте. Здесь также неизвестно, каким образом gp33 и gp34 вытесняют gp28 (или некоторое остаточное количество бактериального фактора σ^{43}), *но как только они связываются с базовым ферментом, он приобретает способность инициировать транскрипцию только на промоторах поздних генов.*

Последовательная замена сигма-фактора имеет двоякое значение. При каждой замене этой субъединицы РНК-полимераза приобретает способность узнавать новый класс генов *и* теряет способность транскрибировать предыдущие гены. Таким образом, эти переключения приводят к системным изменениям в активности РНК-полимеразы. Вероятно, весь (или практически весь) минимальный фермент клетки связывается с тем сигма-фактором, который должен функционировать в данный момент, причем это взаимодействие необратимо.

11.19 Споруляция у бактерий контролируется сигма-факторами

Ключевые понятия:

- В ходе споруляции бактериальная клетка делится на материнскую часть и внутреннюю спору, высвобождающуюся в результате лизиса материнской клетки.
- При споруляции будущая спора и материнская клетка проходят через несколько стадий развития, каждый раз синтезируя новый сигма-фактор (замещающий предыдущий сигма-фактор).
- Скоординированное взаимодействие между спорой и материнской клеткой определяет время замены сигма-факторов.

Возможно, самым ярким примером смены сигма-факторов является **споруляция** — процесс, свойственный ряду бактерий, и альтернативный обычному жизненному циклу бактериальной клетки. В конце **вегетативной фазы** жизненного цикла бактериальной культуры логарифмический рост числа клеток прекращается (поскольку запас питательных веществ среды начинает истощаться). Как показано на рис. 11.41, остановка роста культуры запускает споруляцию. При этом происходят репликация ДНК и сегрегация генома, который локализуется на одном конце клетки и, в конечном счете, покрывается прочной споровой оболочкой. При формировании перегородки образуются два независимых компартмента: материнская клетка и предспора. Процесс спорообразования начинается с того, что две хромосомы, образовавшиеся в результате репликации бактериальной хромосомы, расходятся к противоположным полюсам клетки. Растущая перегородка частично захватывает одну из хромосом в предспору, после чего транслоказа (SpoI-IIE) переносит оставшуюся часть хромосомы в предспору.

Споруляция занимает около восьми часов. Этот процесс может быть рассмотрен как простейший



Споруляция представляет собой разделение вегетативной бактериальной клетки на материнскую клетку, которая затем лизируется, и спору, которая впоследствии высвобождается

пример дифференцировки, в которой родительская клетка (растущая бактерия) дает начало двум разным дочерним клеткам с разной дальнейшей судьбой: материнская клетка в конечном счете лизируется, а высвобождающаяся спора имеет структуру, отличающуюся от структуры исходной бактериальной клетки.

Споруляция характеризуется значительными изменениями биосинтетических процессов в бактерии при участии множества бактериальных генов. Контроль этого процесса осуществляется преимущественно на уровне транскрипции. Транскрипция ряда генов, функционировавших на вегетативной фазе роста, выключается, хотя большинство генов продолжает работать. Кроме того, экспрессия ряда генов свойственна только для периода споруляции. По окончании споруляции около 40% бактериальной мРНК оказывается специфично именно для этой стадии.

В спорулирующих клетках активируются новые формы РНК-полимеразы; в своем составе они содержат такой же минимальный фермент, как и РНК-полимераза вегетативных клеток, но субъединица σ^{43} заменена на другие белки. Изменения в транскрипционной специфичности представлены на РИС. 11.42. Основной принцип этих изменений заключается в том, что в каждом компартменте функционирующий на тот момент сигма-фактор замещается новым фактором, инициирующим транскрипцию новой группы генов. Взаимодействие материнской клетки и предспоры проявляется в координировании времени переключения с одного сигма-фактора на другой.

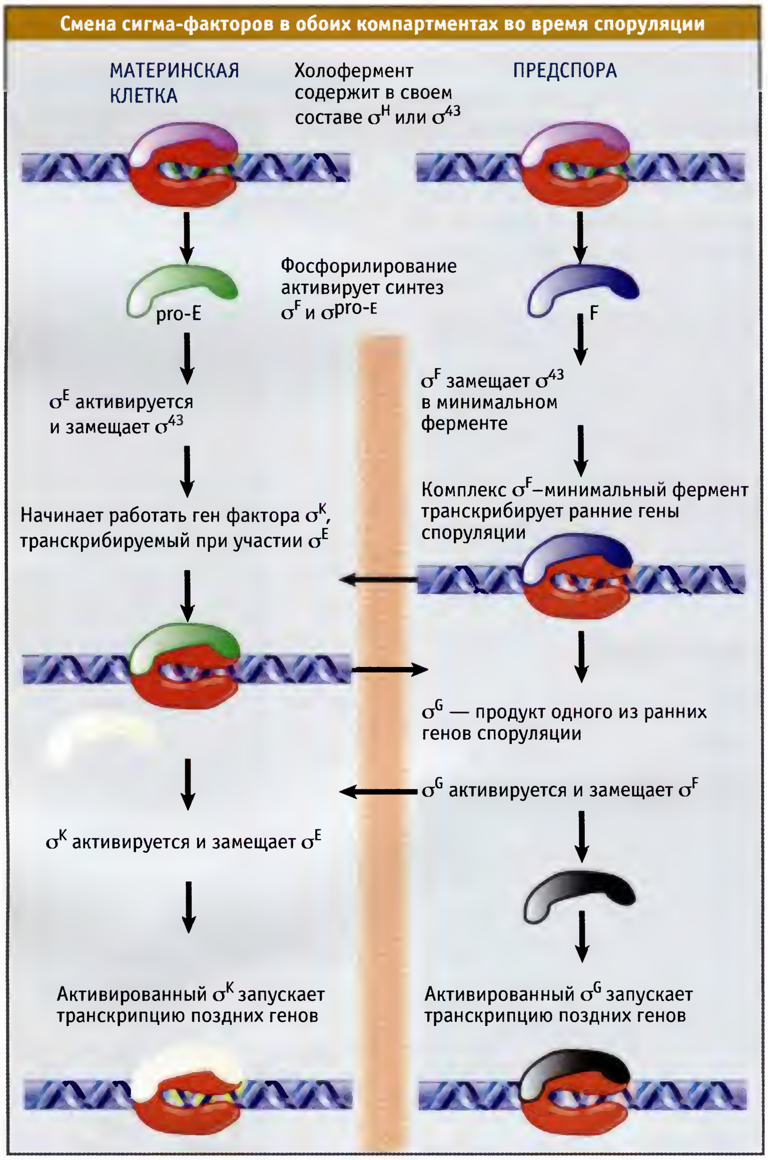


РИС. 11.42 Специфичность инициирующей активности РНК-полимеразы во время споруляции контролируется с помощью смены сигма-факторов. Каскады реакций в материнской клетке (слева) и предспоре (справа) координируются посредством сигналов, передающихся через септу (показаны горизонтальными стрелками)

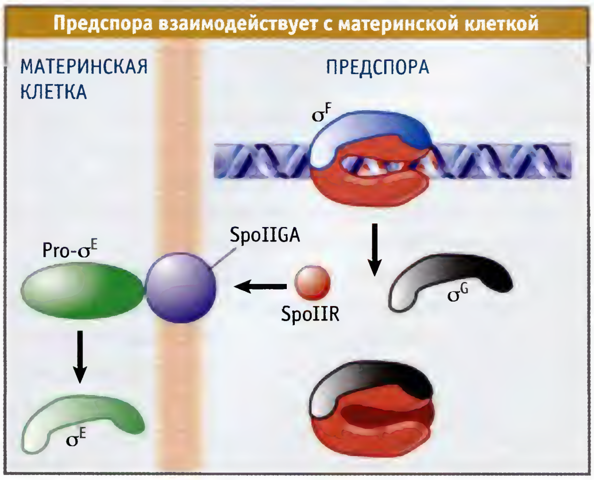


РИС. 11.43 Фактор σ^F запускает процесс синтеза следующего сигма-фактора в предспоре (σ^G) и активирует SpoIIR, который, в свою очередь, активирует SpoIIGA, участвующий в превращении σ^{pro-E} в σ^E

Каскад реакций, составляющих процесс спорообразования, инициируется после того, как в ответ на условия среды запускается **цепь реакций форфорилирования**, в ходе которых фосфатная группа как бы передается через серию кросс-фосфорилирований белку SpoOA. (Вовлечение в этот процесс продуктов сразу нескольких генов отражает важность контроля над запуском процесса спорулирования.) SpoOA — регуляторный белок, активность которого зависит от его фосфорилирования. В фосфорилированной форме он активирует транскрипцию двух оперонов, каждый из которых транскрибируется своей формой РНК-полимеразы. Под влиянием фосфорилированного SpoOA, используя основной фактор σ^{43}, фермент транскрибирует ген, кодирующий фактор σ^F. Используя дополнительный фактор σ^H, фермент транкскрибирует ген, кодирующий фактор pro-σ^E. Оба новых сигма-фактора продуцируются еще до формирования перегородки, однако остаются активными и после ее образования.

Фактор σ^F активируется в предспоре первым. До нужного момента он ингибирован связью с так называемым анти-сигма-фактором; однако в предспоре анти-сигма-фактор удаляется. Реакция контролируется серией событий фосфорилирования/дефосфорилирования. Решающую роль в этом процессе играет фосфатаза (SpoIIE) — мембранный белок, аккумулирующийся на одном из полюсов клетки, вследствие чего фосфатазный домен этого белка концентрируется главным образом в предспоре. SpoIIE дефосфорилирует (и тем самым активирует) SpoIIAA, который, в свою очередь, вытесняет анти-сигма-фактор SpoIIAB из комплекса SpoIIAB-σ^F, и фактор σ^F активируется в результате своего высвобождения из комплекса.

Активация σ^F обозначает начало процесса спорулирования. Под влиянием фактора σ^F РНК-полимераза транскрибирует первую группу генов спорулирования (вместо работавших ранее вегетативных генов). Вероятно, реакция замены сигма-фактора затрагивает только часть РНК-полимераз клетки, поскольку σ^F продуцируется в небольших количествах. Некоторое количество фермента вегетативной фазы остается в клетке и во время споруляции. Замещенный σ^{43} не разрушается и может быть выделен из экстрактов спорулирующих клеток.

Два регуляторных события, следующие за активацией σ^F, детально показаны на РИС. 11.43. В предспоре функционирует другой фактор — σ^G, представляющий собой продукт одного из ранних генов споруляции. Фактор σ^G инициирует транскрипцию поздних споруляционных генов. Продукт еще одного раннего гена отвечает за взаимодействие предспоры с компартментом материнской клетки. Фактор σ^F активирует SpoIIR, который секретируется из предспоры. Активируя мембрансвязанный белок SpoIIGA, он участвует в превращении неактивного предшественника pro-σ^E в активный фактор σ^E в материнской клетке. (Любой σ^E, продуцирующийся в предспоре, деградирует посредством специальных механизмов.)

Каскад реакций продолжается, и σ^E в свою очередь заменяется фактором σ^K. (В действительности, продукция σ^K представляет собой чрезвычайно сложный процесс, поскольку сначала ген этого фактора должен образоваться в результате рекомбинации!) Фактор σ^K синтезируется в виде неактивного предшественника (pro-σ^K), который активируется в результате действия протеазы. Активированный σ^K замещает σ^E и обеспечивает транскрипцию поздних генов в материнской клетке. Протекание всех этих событий в двух компартментах координируется серией сигналов. Активность σ^E в материнской клетке является условием активации σ^G в проспоре; в свою очередь, активность σ^G требуется для генерации сигнала, передающегося через перегородку в материнскую клетку и активирующего σ^K.

Таким образом, споруляция контролируется двумя каскадами реакций, в ходе которых сигма-факторы в каждом компартменте активируются последовательно, обеспечивая транскрипцию определенных групп генов. На РИС. 11.44 изображено, как два описанных каскада ре-



РИС. 11.44 Перекрестная регуляция процесса координирует время споруляционных процессов в материнской клетке и предспоре

акций связаны посредством передачи сигналов из одного компартмента в другой. При активации новых сигма-факторов старые вытесняются, что ведет к выключению транскрипции одних генов и включению других. Встраивание каждого нового сигма-фактора в РНК-полимеразу определяет, какая группа генов будет экспрессироваться, а количество доступного фактора влияет на уровень экспрессии генов. Одновременно могут быть активированы несколько сигма-факторов, причем специфичности некоторых из них перекрываются. Неизвестно, посредством каких механизмов каждый сигма-фактор замещает своего предшественника.

11.20 Бактериальная РНК-полимераза завершает транскрипцию в определенных сайтах

Ключевые понятия:

- Условием окончания транскрипции является узнавание терминаторной последовательности, определяющей формирование шпилечной структуры в РНК-продукте.

Как только РНК-полимераза инициирует транскрипцию, фермент начинает продвигаться вдоль матрицы, синтезируя РНК до тех пор, пока не достигнет **терминирующей** последовательности (терминатора). Здесь РНК-полимераза перестает включать нуклеотиды в растущую цепь РНК, освобождает образовавшийся продукт и диссоциирует от ДНК. При терминации все водородные связи, стабилизирующие РНК-ДНК гибрид, должны быть разрушены, только после этого происходит восстановление двухцепочечной структуры ДНК.

Для молекулы РНК, синтезируемой в живой клетке, сложно определить реальную точку терминации транскрипции. Возможно, 3′-конец молекулы образуется в результате *расщепления* первичного транскрипта и, следовательно, не соответствует реальному сайту, в котором завершила работу РНК-полимераза.

Наилучшим образом сайт терминации можно охарактеризовать, используя систему транскрипции *in vitro*. Способность фермента терминировать синтез РНК во многом зависит от условий реакции (например, таких, как ионная сила). Поэтому терминация, наблюдаемая *in vitro* в какой-либо точке, еще не означает, что данный сайт является терминатором *in vivo*. Об истинном положении терминирующей последовательности можно с уверенностью говорить лишь в том случае, если РНК имеет одинаковый 3′-конец и для *in vitro*, и для *in vivo* систем.

На рис. 11.45 описаны два общих для всех бактериальных терминаторов свойства:

- Бактериальные и фаговые терминаторы были обнаружены как последовательности, необходимые для осуществления реакции терминации (*in vivo* или *in vitro*). Анализ последовательностей прокариотических терминаторов не обнаружил никакого сходства для регионов, находящихся ниже точки окончания синтеза РНК. Это указывает на то, что участок, ответственный за терминацию, расположен в пределах *последовательности, транскрибируемой* РНК-полимеразой. Действительно, терминация зависит от свойств определенных участков молекулы РНК, синтезируемой РНК-полимеразой.
- Многие терминаторы для осуществления своей функции требуют формирования во вторичной структуре синтезируемой РНК специальной шпильки. *Это лишний раз свидетельствует в пользу того, что терминация определяется свойствами РНК-продукта.*

Терминаторы сильно различаются по эффективности терминации. На некоторых терминаторах событию терминации могут препятствовать дополнительные регуляторные белки, взаимодействующие с РНК-полимеразой. В результате такой **антитерминации** фермент продолжает синтезировать РНК за пределы терминатора. Это явление называют **прочитыванием** терминатора (этот же термин используется в разд. 9.14 «Супрессорные тРНК могут конкурировать с тРНК дикого типа за связывание с кодонами»).

Терминацию необходимо рассматривать не просто как механизм образования 3′-конца молекулы РНК, а как способ контроля экспрессии генов. Не только связывание РНК-полимеразы с ДНК (инициация), но и диссоциация фермента от матрицы (терминация) определенным образом контролируются клеткой. Можно обнаружить интересные аналогии между системами, участвующими в инициации и терминации транскрипции. В обоих

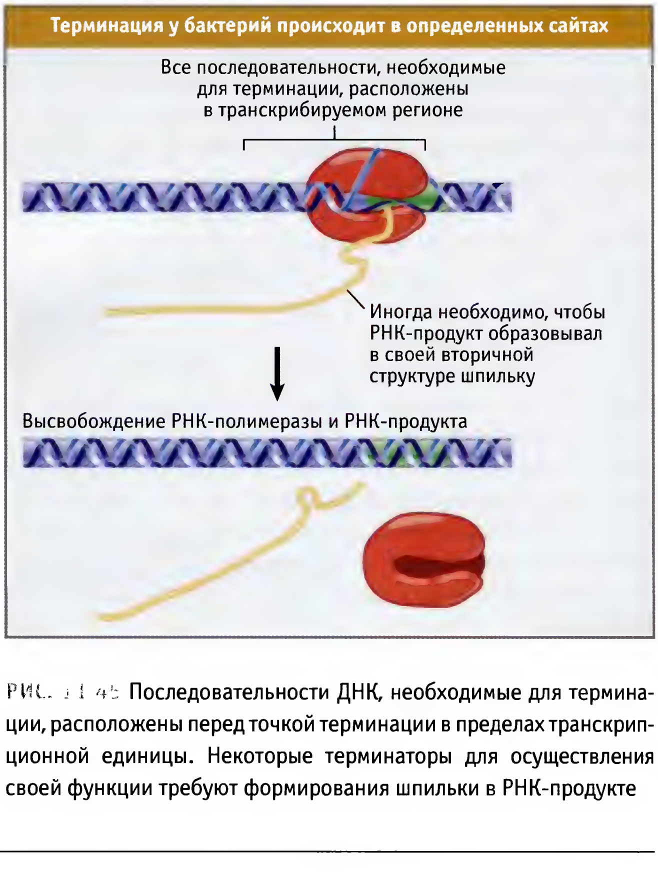


РИС. 11.45 Последовательности ДНК, необходимые для терминации, расположены перед точкой терминации в пределах транскрипционной единицы. Некоторые терминаторы для осуществления своей функции требуют формирования шпильки в РНК-продукте

случаях должен происходить разрыв водородных связей (первоначальное плавление ДНК в случае инициации, и диссоциация гибрида РНК-ДНК при терминации). Также в обоих случаях требуются дополнительные белковые факторы, взаимодействующие с минимальным ферментом. Можно сказать, что эти процессы осуществляются альтернативными формами полимеразы. В то время как инициация зависит исключительно от взаимодействия РНК-полимеразы с двухцепочечной структурой ДНК, на терминацию влияют как сигналы самого транскрипта, так и особенности последовательности ДНК, распознаваемые полимеразой и дополнительными регуляторными факторами.

11.21 Два типа терминаторов *E. coli*

Ключевые понятия:

- Независимые терминаторы состоят из G-C-богатой шпильки РНК и последующего за ней U-богатого региона, на котором происходит терминация.



РИС. 11.46 Р-независимые терминаторы содержат в своем составе палиндромы, формирующие шпильки, длина которых варьирует от 7 до 20 пн. В основании структуры стебель–петля расположен G-C-богатый регион, за которым находится участок, состоящий из остатков U

В зависимости от того, требуются ли РНК-полимеразе для осуществления терминации *in vitro* какие-либо дополнительные факторы, терминаторы *E.coli* делят на две группы:

- Минимальный фермент может терминировать транскрипцию в определенных сайтах *in vitro* в отсутствие каких бы то ни было дополнительных факторов. Такие сайты называют **ρ-независимыми терминаторами**.
- **ρ-Зависимые терминаторы** функционируют *in vitro* только в присутствии **ρ-фактора**. Изучение мутаций по этому фактору показало, что фактор ρ участвует в терминации транскрипции и *in vivo*.

Независимые терминаторы имеют две структурные особенности, изображенные на РИС. 11.46: короткий инвертированный повтор, способный формировать двухцепочечный участок (шпильку) во вторичной структуре, и участок, расположенный в самом конце транскрипционной единицы, богатый остатками U. Обе эти особенности являются необходимым условием терминации. Шпилька обычно содержит G-C-богатый регион, расположенный у основания стебля. Расстояние от шпильки до U-богатого участка составляет, как правило, от семи до девяти оснований. В геноме *E.coli* описано порядка 1100 последовательностей, удовлетворяющих этим критериям. Это означает, что около половины всех генов обладают ρ-независимыми терминаторами.

Предотвращающие терминацию точечные мутации обычно затрагивают участок, соответствующий двухцепочечному региону шпилечной структуры. Какой эффект шпилька оказывает на транскрипцию? Вероятно, все шпильки, образуемые РНК продуктом, замедляют (или даже останавливают) работу РНК-полимеразы.

Возникающая пауза способствует терминации транскрипции. Приостановка транскрипции происходит и в сайтах, по своей структуре напоминающих терминаторы, но характеризующихся увеличенным расстоянием между шпилькой и U-богатым участком (как правило, 10–12 оснований). Если сайт, в котором произошла остановка транскрипции, не является терминатором, фермент начинает двигаться снова, продолжая транскрипцию. Продолжительность паузы варьирует, однако на типичных терминаторах она длится около 60 секунд.

Расположенный за шпилькой U-богатый регион дестабилизирует РНК-ДНК гибрид в тот момент, когда РНК-полимераза делает паузу, «столкнувшись» со шпилькой. Сила связывания между основаниями rU-dA в гетеродуплексе РНК-ДНК, как правило, довольно слабая; для разрушения связей между цепями в таком гибриде требуется меньшее количество энергии. Когда полимераза делает паузу, гетеродуплекс разрушается начиная с концевого региона, богатого слабыми связями rU-dA. Терминация может произойти на любом из нескольких остатков U, расположенных ближе к концу U-участка, при этом фермент как бы «запинается» на этом участке. Последовательность из оснований U соответствует A-T-богатой области в ДНК, из чего следует, что A-T-богатые участки играют важную роль как при терминации, так и при инициации.

На эффективность терминации оказывает влияние как последовательность участка, образующего шпильку, так и длина U-богатого региона. Эффективность терминации *in vitro*, однако, варьирует от 2% до 90% и не обнаруживает при этом какой-либо заметной корреляции с составом шпильки или числом уридинов в U-богатом регионе. Это означает, что шпилька и U-участок явля-

ются необходимыми, но недостаточными условиями терминации, и на работу РНК-полимеразы оказывает влияние ряд дополнительных факторов. В частности, последовательности, окружающие терминатор, влияют на эффективность терминации.

Гораздо меньше известно о сигналах и дополнительных факторах, влияющих на терминацию эукариотических полимераз. Каждый класс эукариотических полимераз использует свой собственный механизм (см. гл. 26 «Сплайсинг и процессинг РНК»).

11.22 Как работает ρ-фактор?

Ключевые понятия:

- **ρ-фактор представляет собой терминирующий белок, который связывает насцентную цепь РНК в сайте *rut*, движется вдоль нее и отсоединяет от РНК-ДНК гибрида при достижении РНК-полимеразы.**

Фактор ρ — белок *E. coli*, функционирующий только во время терминации. Он действует на ρ-зависимые терминаторы, составляющие примерно половину всех терминаторов *E.coli.*

На РИС. 11.47 показано, как функционирует ρ-фактор. Он связывается с последовательностью РНК, расположенной внутри транскрипционной единицы перед сайтом терминации. Этот сайт называют сайтом ***rut*** (аббревиатура от ρ utilization). Затем ρ-фактор начинает двигаться вдоль цепи РНК до тех пор, пока не нагонит РНК-полимеразу. Когда полимераза достигает сайта терминации, ρ-фактор взаимодействует с находящимся внутри фермента гетеродуплексом РНК-ДНК, вызывая высвобождение РНК из комплекса. Остановка РНК-полимеразы на терминаторе дает возможность ρ-фактору провзаимодействовать с гибридной структурой РНК-ДНК, что является ключевым моментом терминации.

На этом примере можно рассмотреть важный принцип работы ДНК-связывающих белков: зная точный сайт последовательности ДНК, в котором тот или иной белок оказывает свой эффект, мы не можем утверждать, что этот же сайт участвует в первоначальном связывании белка с ДНК. Эти участок «срабатывания» и участок первичного связывания могут быть удалены друг от друга и не обнаруживать между собой никакой видимой связи. На практике сайты *rut* в разных транскрипционных единицах обнаруживаются на разных расстояниях от сайтов терминации. Похожая ситуация наблюдается и в случае с антитерминирующими факторами (см. разд. 11.24 «Антитерминация опосредована сайтами, отличными от терминаторов»).

Общей чертой всех сайтов *rut* является большое содержание остатков C и малое количество остатков G в последовательности сайта, вследствие чего можно предположить отсутствие шпилек в их вторичной структуре. Пример такой последовательности приведен на РИС. 11.48. Видно, что основание C встречается чаще других (41%), в то время как G — реже всего (14%). Длина сайтов *rut*

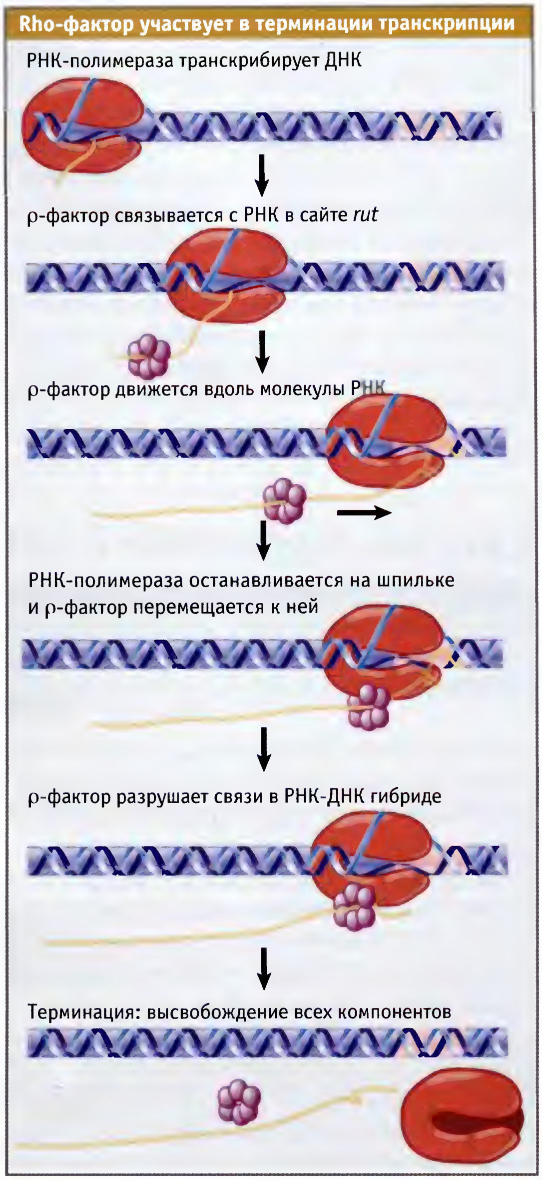


РИС. 11.47 ρ-фактор связывается с РНК в сайте *rut* и движется вдоль молекулы РНК до тех пор, пока не достигнет РНК-ДНК гибрида внутри РНК-полимеразы. Здесь он освобождает РНК из связи с ДНК

варьирует. Как правило, эффективность сайта *rut* возрастает с увеличением длины C-богатого участка.

ρ-Фактор является членом семейства гексамерных АТФ-зависимых хеликаз. Субъединица этого белка имеет РНК-связывающий домен и АТФ-гидролизующий домен. РНК-связывающие домены шести субъединиц образуют канал, через который проходит нуклеиновая кислота. На РИС. 11.49 приведена структура ρ-фактора, объясняющая, как именно он функционирует. Фактор связывает молекулу РНК образующими кольцо N-концевыми доменами, после чего проталкивает 5′-конец связанного участка транскрипта в свою центральную область, где тот связывается с вторичным РНК-связывающим участком C-концевого домена. Первоначально ρ-фактор имеет форму

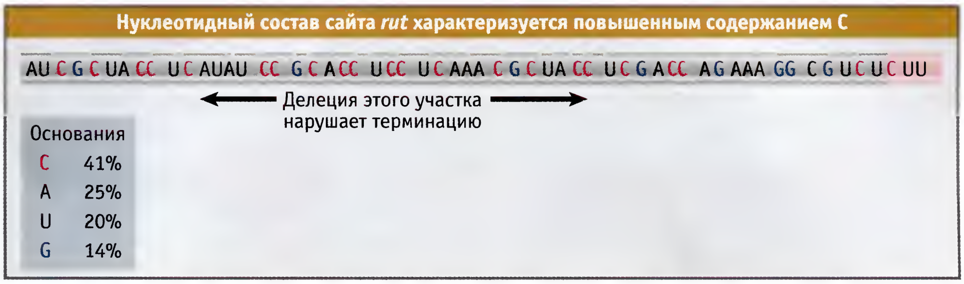


РИС. 11.48 Сайт *rut* расположен перед точкой терминации. Он представлен последовательностью, богатой остатками С и бедной остатками G. Эта последовательность соответствует 3′-концу молекулы РНК

разорванного кольца, однако после связывания с РНК превращается в замкнутое кольцо.

После связывания с сайтом *rut* ρ-фактор, используя свою хеликазную активность и энергию гидролиза АТФ, начинает двигаться вдоль цепи РНК до тех пор, пока не достигнет гибридного участка РНК-ДНК, находящегося внутри РНК-полимеразы. Здесь фактор расплетает дуплексную структуру гибрида. Мы не знаем, является это событие достаточным условием для высвобождения транскрипта или же ρ-фактору необходимо провзаимодействовать с РНК-полимеразой. ρ-Фактор служит дополнительным регуляторным фактором для РНК-полимеразы; как правило, его максимальная активность проявляется, когда концентрация фактора составляет примерно 10% от количества РНК-полимераз.

ρ-фактор перемещается вдоль цепи РНК от сайта *rut* до точки терминации. Для этого необходимо, чтобы ρ-фактор двигался быстрее, чем РНК-полимераза. Достигая терминатора, фермент делает паузу, и терминация происходит только в том случае, если ρ-фактор догоняет полимеразу именно в этот момент. Как уже было сказано, пауза в транскрипции важна для функционирования как ρ-зависимых, так и ρ-независимых терминаторов, поскольку она обеспечивает время для осуществления дополнительных терминационных событий.

Идея о том, что ρ-фактор передвигается вдоль РНК, позволила высказать важное предположение о взаимосвязи процессов транскрипции и трансляции. Для того чтобы ρ-фактор мог выполнить свою работу, ему необходимо, во-первых, провзаимодействовать с сайтом связывания на РНК и, во-вторых, продвигаться вдоль молекулы транскрипта. Интересно, что у бактерий транслирующие мРНК рибосомы препятствуют выполнению каждого из этих условий. Поэтому достигнет ли фактор терминатора (на котором сделала остановку РНК-полимераза), во многом зависит от событий, происходящих во время трансляции еще синтезируемого транскрипта.

С помощью этой модели можно объяснить явления, которое длительное время вызывало недоумение. У бактерий нонсенс-мутация в одном из генов может нарушать экспрессию расположенных за ним генов, входящих в состав той же транскрипционной единицы. Данный эффект был назван **полярностью**. Основная причина

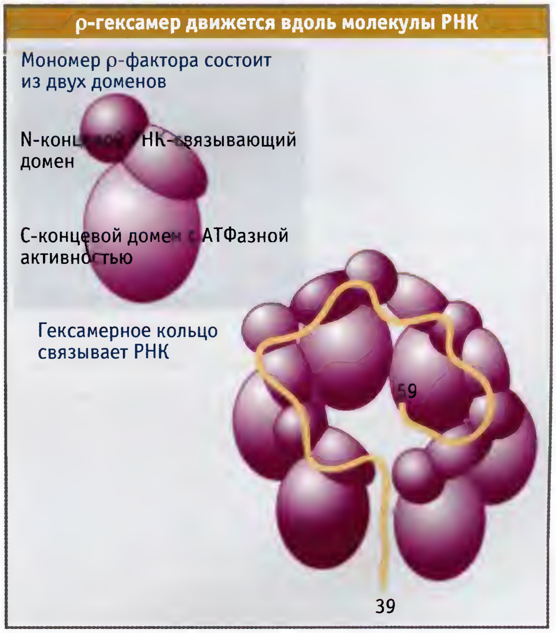


РИС. 11.49 В структуре ρ-фактора различают N-концевой РНК-связывающий домен и С-концевой домен с АТФазной активностью. Гескамер в форме разомкнутого кольца связывает РНК вдоль внешней поверхности N-концевых доменов. 5′-конец РНК связывается со вторичным сайтом внутри гексамерной структуры фактора

данного явления — отсутствие мРНК, соответствующей дистальной части транскрипционной единицы. Предположим, что *внутри* транскрипционной единицы существуют ρ-зависимые терминаторы, расположенные перед *основным* терминатором. Дальнейшие события можно объяснить механизмами, показанными на РИС. 11.50. Обычно такие («внутренние») терминаторы не работают, так как из-за присутствия рибосом фактор на этих терминаторах не взаимодействует с РНК-полимеразой. Нонсенс-мутации приводят к диссоциации рибосом от мРНК, и поэтому у ρ-фактора появляется возможность свободно передвигаться вдоль мРНК. В результате он может провзаимодействовать с полимеразой, задержавшей-

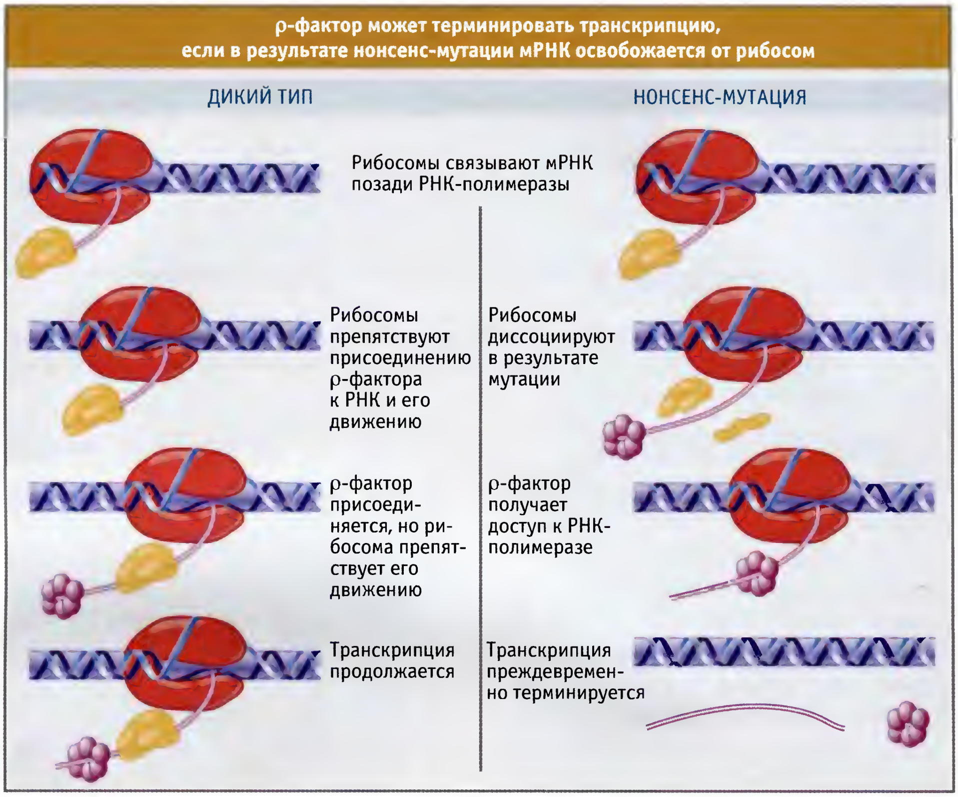


РИС. 11.50 Фактор ρ может выступать в роли связующего звена между процессами транскрипции и трансляции в том случае, когда ρ-зависимые терминаторы располагаются сразу после сайта с нонсенс-мутацией

ся на «внутреннем» терминаторе. Это, в свою очередь, приведет к освобождению РНК-полимеразы и невозможности экспрессии дистальной части транскрипционной единицы. (Для чего существуют внутренние терминаторы? Возможно, они просто представляют собой последовательности, своей структурой напоминающие обычные терминаторы.) Некоторые стабильные РНК защищены от действия эффекта полярности благодаря своей жесткой вторичной структуре, которая препятствует связыванию и движению ρ-фактора.

Мутации в гене *rho* в разной степени сказываются на процессе терминации. Общий эффект таких мутаций заключается в нарушении терминации, однако степень повреждения, если ее оценивать по проценту случаев прочитывания терминатора *in vivo*, зависит от конкретного терминатора. Похожим образом варьирует и необходимость участия в терминации *in vitro* самого ρ-фактора. В то время как некоторые (ρ-зависимые) терминаторы требуют для своего функционирования высоких концентраций фактора, другие функционируют при относительно низких концентрациях. Если для работы различных терминаторов необходимо разное количество факторов, это означает, что оставшаяся в мутантных клетках активность ρ-фактора может обеспечить работу некоторых терминаторов, для которых требуется низкая концентрация фактора (большинство мутаций по гену ρ являются «подтекающими» (с неполным блокированием функции)).

Некоторые мутации в гене ρ-фактора могут быть супрессированы мутациями, возникающими в других генах. Этот факт дает возможность выявить белки, контактирующие с ρ. Участие β-субъединицы РНК-полимеразы во взаимодействии с ρ можно подтвердить благодаря существованию мутаций двух типов. Во-первых, мутации в гене *rpoB* могут ослаблять терминацию в ρ-зависимых сайтах. Во-вторых, введение мутаций в ген *rpoB* может восстановить способность терминировать транскрипцию в ρ-зависимых сайтах у бактерий, дефектных по этому фактору. Тем не менее пока неясно, какую именно роль играет взаимодействие β-субъединицы и ρ-фактора.

11.23 Антитерминация — специальный регуляторный механизм

Ключевые понятия:

- Терминация может быть предотвращена воздействием антитерминирующего белка на РНК-полимеразу, что вызывает прочитывание ферментом специфического терминатора или терминаторов.
- Фаг лямбда имеет два антитерминирующих белка, pN и pQ, действующих на разные транскрипционные единицы.

Антитерминация используется в качестве контрольного механизма как в регуляторных системах фагов, так и в бактериальных оперонах. На РИС. 11.51 показано, как антитерминация контролирует способность фермента прочитывать терминатор и транскрибировать расположенные за ним гены. В приведенном на рисунке примере РНК-полимераза должна останавливать транскрипцию в конце региона 1; однако антитерминация позволяет ферменту продолжать транскрипцию и считывать регион 2. Поскольку промотор в обеих ситуациях остается одним и тем же, возникают молекулы РНК с одинаковыми 5′-концами; однако во втором случае 3′-конец транскрипта представлен новой, более протяженной, последовательностью.

Антитерминация была открыта при изучении фаговых инфекций бактерий. Общим свойством фаговых инфекций является то, что только небольшая часть генов («ранних» фаговых генов) транскрибируется при помощи бактериальной РНК-полимеразы. При этом один (или несколько) из ранних фаговых белков необходим для транскрипции последующего набора генов (см. разд. 14.4 «Литический каскад контролируется двумя способами»). Одним из таких регуляторных белков является **антитерминирующий белок**. На РИС. 11 показано, как, делая возможным прочитывание РНК-полимеразой терминатора, антитерминирующий белок способствует образованию транскрипта большей длины. В отсутствие теминирующего белка фермент останавливает свою работу в области терминатора (верхняя часть рисунка). В присутствии антитерминирующего белка РНК-полимераза продолжает транскрибировать участки, расположенные за терминатором (средняя часть рисунка).

Наиболее изученным примером антитерминации является регуляция транскрипции при развитии фага лямбда (на котором это явление и было открыто). Антитерминация у фага лямбда используется для регуляции двух стадий экспрессии фаговых генов. Продуцируемые на каждой из этих стадий антитерминирующие белки специфичны к определенному набору транскрипционных единиц (см. нижнюю часть рис. 11.52).

Два фаговых гена, транскрибируемых бактериальной РНК-полимеразой сразу после проникновения генома фага в клетку, называют **немедленноранними** (immediate early) генами. Переключение на экспрессию следующей группы генов контролируется нарушением терминации при транскрипции немедленноранних генов. Гены, экспрессирующиеся на следующей стадии, называют **задержанноранними** (delayed early). (Общие механизмы регуляции развития фага лямбда рассмотрены в гл. 14 «Фаговые стратегии».)

Регуляторный ген, продукт которого контролирует переключение транскрипции с немедленноранних генов на задержанноранние, был идентифицирован благодаря мутациям, нарушающим это переключение. Мутанты фага лямбда по гену *N* способны транскрибировать *только* немедленноранние гены, и поэтому инфекционный процесс останавливается на этой стадии.

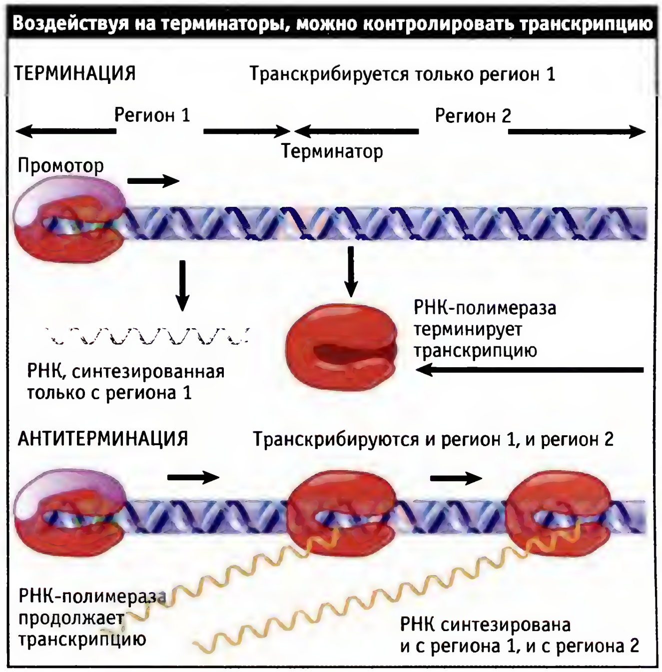
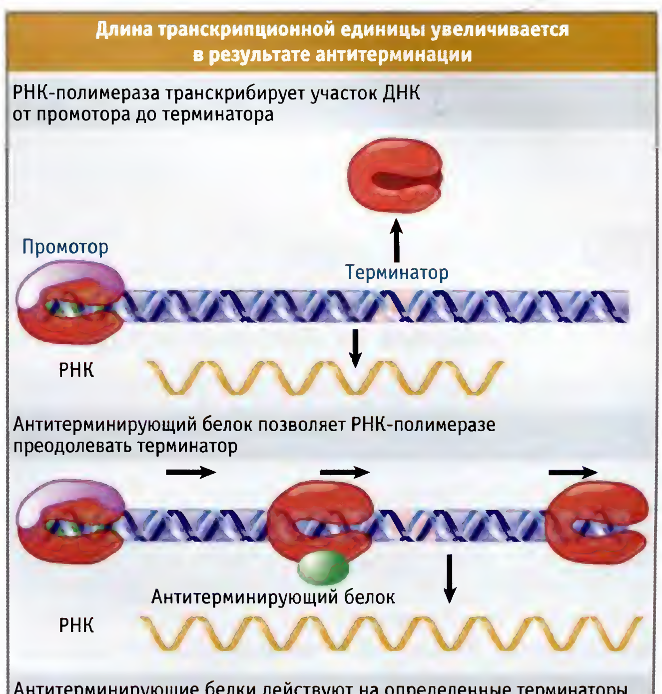


РИС. 11.51 Антитерминация определяет, прекратит ли РНК-полимераза транскрипцию, достигнув терминатора, или же, прочитав его, продолжит считывать регионы, расположенные за ним



Транскрипционная единица	Промотор	Терминатор	Антитерминирующий белок
Немедленно ранняя	P_L	t_L	pN
Немедленно ранняя	P_{R1}	t_{R1}	pN
Поздняя	$P_{R'}$	$t_{R'}$	pQ

РИС. 11.52 Антитерминирующий белок действует на РНК-полимеразу, позволяя ей преодолевать некоторые терминаторы

Определены две транскрипционные единицы немедленноранних генов (транскрибируемые с промоторов P_L и P_R). Транскрипция, осуществляемая РНК-полимеразой *E. coli,* останавливается в конце транскрипционных единиц (на терминаторах t_{L1} и t_{R1} соответственно). Оба этих терминатора являются ρ-зависимыми (благодаря функционированию именно этих терминаторов фактор ρ и был обнаружен). Ситуация меняется благодаря экспрессии гена *N*. Продуктом этого гена является антитерминирующий белок, позволяющий РНК-полимеразе прочитывать t_{L1} и t_{R1}. В результате транскрибируются задержанно ранние гены, расположенные за этими терминаторами.

Как и у других фагов, в случае фага лямбда для экспрессии поздних генов (кодирующих компоненты фаговой частицы) необходимы дополнительные регуляторные механизмы. Такой контроль осуществляется геном *Q*, относящимся к задержанноранним генам. Продукт этого гена (белок pQ) является еще одним антитерминирующим фактором, позволяющим РНК-полимеразе (инициирующей транскрипцию на позднем промоторе $P_{R'}$) преодолевать терминатор, расположенный между этим промотором и поздними генами.

Разная специфичность белков pN и pQ позволяет сформулировать общий принцип: *РНК-полимераза взаимодействует с транскрипционными единицами таким образом, чтобы антитерминирующие факторы могли избирательно отключать синтез только определенных транскриптов.* Возможно, что терминация при этом контролируется с той же степенью точности, что и инициация.

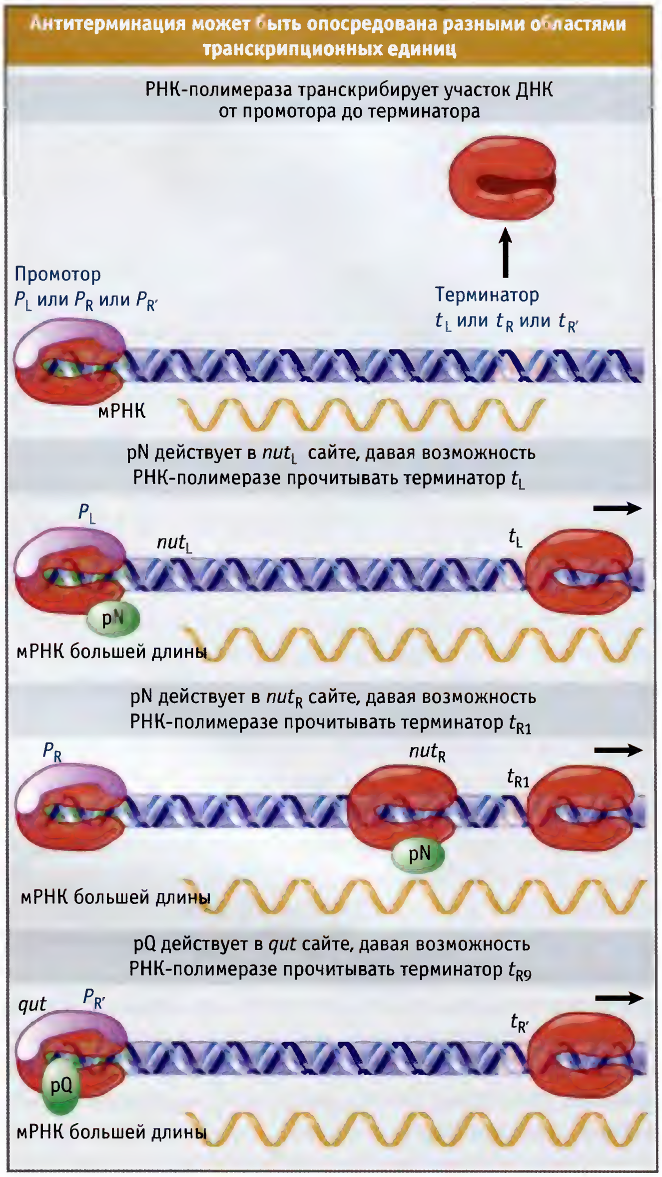


РИС. 11.53 РНК-полимераза хозяина транскрибирует гены фага лямбда и терминирует транскрипцию в *t*-сайтах. Белок pN позволяет ей преодолевать терминаторы в транскрипционных единицах L и R1; белок pQ — соответственно терминатор R′. Сайты, в которых действуют белки pN (*nut*-сайты) и pQ (*qut*-сайты) расположены в разных областях соответствующих транскрипционных единиц

11.24 Антитерминация опосредована сайтами, отличными от терминаторов

Ключевые понятия:

- Сайт, который связывает антитерминирующий белок, расположен в транскрипционной единице выше сайта терминации.
- Местоположение сайта антитерминации в разных случаях варьирует, — он может располагаться в кодирующем регионе или даже в промоторе.

Какие сайты определяют специфичность антитерминации? Антитерминирующее действие белка pN высокоспецифично, однако *антитерминационные события не определяются терминаторами t_{L1} и t_{R1}; сайты, распознавание которых ведет к антитерминации, расположены в транскрипционной единице перед терминаторами. Другими словами, сайт, необходимый для антитерминации не совпадает с сайтом терминации.* На РИС. 11.53 изображено местоположение сайтов, участвующих в антитерминации у фага лямбда.

Участок ДНК, распознаваемый белком N, называют *nut* (англ. *N utilization*). Участки, ответственные за антитерминацию при левосторонней и правосторонней транскрипции, обозначают соответственно как *nutL* и *nutR*. Картируя мутации в сайте *nut,* удалось выяснить, что *nutL* расположен между стартовой точкой промотора P_L и началом области, кодирующей ген N. *nutR* находится между концом гена *cro* и терминатором t_{L1}. Это означает, что оба сайта *nut* расположены в совершенно разных областях своих транскрипционных единиц. В

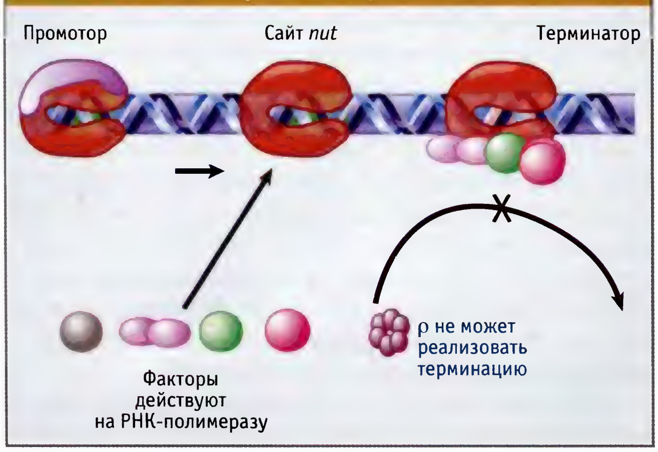


РИС. 11.54 Вспомогательные факторы связывают РНК-полимеразу в тот момент, когда она проходит *nut*-сайт. Когда же полимераза достигает терминатора, они не дают ρ-фактору вызвать терминацию

то время как *nutL* находится рядом с промотором, *nutR* расположен недалеко от терминатора (*qut*-сайт отличается от описанных выше *nut*-сайтов и расположен в промоторе).

Как же осуществляется антитерминация? Когда белок pN распознает сайт *nut*, он должен провзаимодействовать с РНК-полимеразой таким образом, чтобы фермент в дальнейшем уже не распознавал терминатор. Различная локализация *nut* сайтов указывает на то, что это событие не связано напрямую ни с инициацией, ни с терминацией, но может происходить с РНК-полимеразой в момент элонгации в *nut*-сайте. Как показано на РИС. 11.54, после этого полимераза становится настолько «неконтролируемой», что преодолевает терминатор, не обращая на него внимания. (Данный механизм реакции характерен для ρ-зависимых терминаторов, однако любопытно, что белок pN способен супрессировать терминацию и на ρ-независимых терминаторах.)

Является ли способность белка pN узнавать короткую последовательность в транскрипционной единице примером более широко используемого механизма антитерминации? Другие фаги (родственные фагу лямбда) имеют различные гены *N* и различную антитерминаторную специфичность. Область фагового генома, в которой находятся *nut*-сайты, представлена разной последовательностью у разных фагов, поэтому каждый фаг обладает только ему свойственными *nut*-сайтами, специфично узнаваемыми собственным белком pN. Вместе с тем все белки pN должны обладать равной способностью взаимодействовать с транскрипционным аппаратом, осуществляя антитерминирующую функцию, но при этом иметь различную специфичность по отношению к последовательности ДНК, активирующей этот процесс.

11.25 Факторы терминации и антитерминации взаимодействуют с РНК-полимеразой

Ключевые понятия:

- Для взаимодействия pN фага лямбда с РНК-полимеразой необходимо несколько бактериальных белков.
- Эти белки также вовлечены в антитерминацию в опероне *rrn* бактерии-хозяина.
- Взаимодействие антитерминирующего фактора pQ λ-фага с ДНК в области промотора может принимать различные формы.

Процессы терминации и антитерминации тесно связаны между собой и осуществляются при участии бактериальных и фаговых белков, взаимодействующих с РНК-полимеразой. Несколько белков, связанных с процессом терминации, были идентифицированы при выделении мутантов *E.coli*, дефектных по белку pN. Некоторые из этих мутаций располагались в гене *rpoB*. Данный факт говорит в пользу того, что pN (как и ρ-фактор) взаимодействует с β-субъединицей минимального фермента. Другие мутации *E.coli*, блокирующие функцию белка pN, были картированы в локусе *nus*: *nusA, nusB, nusE* и *nusG*. (*nus* является аббревиатурой от *N* *u*tilization *s*ubstance).

Сайт *nut* фага лямбда состоит из двух элементов, названных *boxA* и *boxB*. Последовательности, близкие по составу к *boxA*, были обнаружены и в бактериальных оперонах. *boxA* необходим для связывания бактериальных белков, участвующих в антитерминации фаговых и бактериальных оперонов. *boxB* специфичен для фагового генома, мутации в этом регионе блокируют антитерминирующую способность pN.

Локус *nus* кодирует белки, являющиеся частью транскрипционного аппарата, которые, однако, в составе изолированной РНК-полимеразы не обнаруживаются. Продукты генов *nusA, nusB* и *nusG* участвуют исключительно в процессе терминации транскрипции. *nusE* кодирует рибосомальный белок S10; взаимосвязь между расположением этого белка в 30S субъединице рибосомы и его участием в терминации неясна. Как показано на РИС. 11.55, белок Nus связывает РНК-полимеразу в *nut* сайте.

NusA — основной транскрипционный фактор, повышающий эффективность терминации (вероятно, за счет усиления способности РНК-полимеразы делать паузу на терминаторах и других участках со сложной вторичной структурой; см. ниже). NusB и S10 образуют димер, который специфично связывается с последовательностью *boxA* на РНК-полимеразе. Белок NusG обеспечивает сборку всех факторов Nus в комплекс с РНК-полимеразой.

ρ-Независимые и ρ-зависимые терминаторы задействуют различные Nus-факторы. Для терминации на независимых терминаторах необходим фактор NusA, причем реакция может быть блокирована белком pN. ρ-зависимые терминаторы для осуществления своей функции требуют участия всех факторов Nus, и

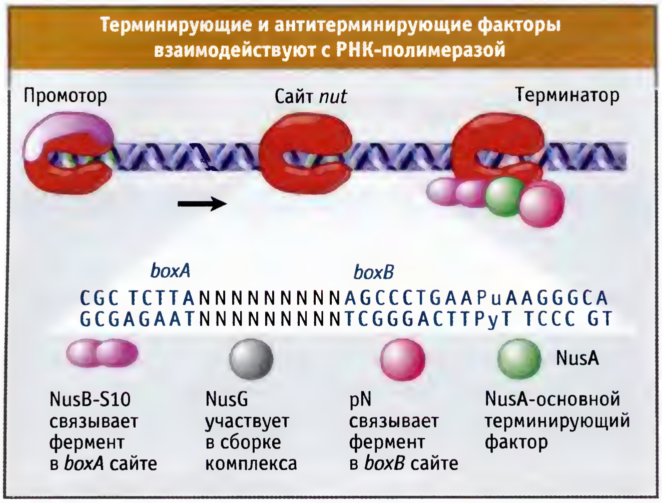


РИС. 11.55 Вспомогательные факторы связываются с РНК-полимеразой в тот момент, когда она проходит через определенные сайты. Сайт *nut* представлен двумя последовательностями. NusB-S10 присоединяется к основному ферменту в тот момент, когда он проходит *boxA*. Белки NusA и pN связываются с полимеразой, когда она проходит *boxB*. В присутствии pN фермент прочитывает терминатор, в результате чего образуется РНК, содержащая в своем составе последовательности как немедленноранних, так и задержанноранних генов

снова одного белка pN достаточно для блокирования терминации. В обоих случаях функция белка pN заключается в ингибировании терминирующей функции NusA. Связывание NusA с pN ингибирует необходимую для осуществления терминации способность NusA взаимодействовать с РНК.

Антитерминация происходит и в оперонах *rrn* (рРНК) *E. coli* и опосредуется теми же функциями локусов *nus*. Лидирующие участки оперонов *rrn* содержат последовательности *boxA*; димеры NusB-S10 распознают эти последовательности и связываются с РНК-полимеразой, когда она проходит через *boxA*. Это взаимодействие изменяет свойства РНК-полимеразы таким образом, что она приобретает способность преодолевать ρ-зависимый терминатор, расположенный внутри транскрипционной единицы.

Димер NusB-S10 не связывает последовательность *boxA* фага лямбда, возможно, это происходит только в присутствии NusA и pN; роль последовательности *boxB* заключается в стабилизации реакции. Вариации последовательностей *boxA* отражают разнообразие регуляторных факторов, требуемых для антитерминации. Когда РНК-полимераза проходит через *nut*-сайт, она модифицируется путем присоединения соответствующего фактора, после чего, достигнув терминатора, уже не останавливается на нем.

Антитерминация у фага лямбда зависит от опосредованного последовательностью транскрипта связывания pN с РНК-полимеразой. Где белок pN узнает *boxB* сайт: в ДНК или в РНК? Этот фактор не связывается с какой-либо из нуклеиновых кислот независимо, pN ассоциирует с транскрипционным комплексом в тот момент, когда базовый фермент проходит через сайт *boxB*. pN имеет отдельные домены для распознавания сайта *boxB* на РНК и для связывания белка NusA. После присоединения к транскрипционному комплексу pN, не теряя связи с минимальным ферментом, становится как бы дополнительной субъединицей, присутствие которой изменяет характер взаимодействия фермента с терминатором. Возможно, что pN в дальнейшем сохраняет связь как с последовательностью *boxB* на РНК (формирующей шпилечную структуру), так и с РНК-полимеразой; в этом случае роль *boxB* будет сведена к связыванию белка pN и увеличению его локальной концентрации.

Действие белка pQ, блокирующего терминацию на поздних стадиях фаговой инфекции и взаимодействующего соответственно с *qut*-сайтами ДНК, описывается несколько иными механизмами. Последовательность *qut* расположена в районе стартовой точки поздней транскрипционной единицы фага. Начальная часть последовательности *qut* находится внутри промотора, тогда как оставшаяся часть расположена в начале транскрибируемого региона. Это означает, что pQ должен обладать способностью специфично распознавать ДНК, а механизм действия данного регуляторного фактора, по крайней мере в момент первичного связывания с транскрипционным комплексом, должен отличаться от такового для белка pN. pQ взаимодействует с холоферментом во время инициации. Это взаимодействие требует участия фактора σ^{70}. Данный факт подтверждает «пролонгированные» свойства конформационных изменений РНК-полимеразы, когда изменения в структуре фермента, возникающие на одной стадии процесса транскрипции, могут влиять на его активность на последующих стадиях.

Основная роль белка pQ заключается в изменении продолжительности паузы, которую РНК-полимераза делает на терминаторах. Взаимодействуя с ферментом, он способствует сокращению паузы на всех терминирующих последовательностях, включая как ρ-зависимые, так и ρ-независимые терминаторы. pQ не влияет на процесс терминации напрямую, однако дает возможность РНК-полимеразе быстрее преодолевать терминаторы, тем самым нарушая способность минимального фермента и/или дополнительных факторов осуществлять терминацию.

Согласно общим представлениям, РНК-полимераза может существовать в разных формах, каждая из которых ответственна за реализацию отдельных стадий процесса транскрипции; активность РНК-полимеразы каждой из этих стадий может быть изменена только путем модификации формы фермента. Замена сигма-факторов может вызвать смену одной формы фермента (ответственной за инициацию) на другую; присутствие факторов Nus может повлиять на свойства ответственной за терминацию формы РНК-полимеразы.

По-видимому, процесс терминации тесно связан с элонгацией. Во время удлинения цепи РНК минималь-

ный фермент часто делает остановки, пауза же на терминирующей последовательности служит предпосылкой для терминации. Под влиянием таких регуляторных факторов, как NusA, продолжительность паузы увеличивается, и эффективность терминации повышается. В то же время, под влиянием белков pN и pQ, пауза сокращается, и эффективность терминации падает. Сайты распознавания и связывания перечисленных факторов найдены только в некоторых транскрипционных единицах, а потому особенности процесса терминации варьируют только для этих транскриптонов.

11.26 Резюме

Транскрипционная единица представляет собой область ДНК, заключенную между промотором, где транскрипция инициируется, и терминатором, где транскрипция обрывается. Одна из цепей ДНК этого региона служит матрицей для синтеза комплементарной цепи РНК. Во время движения транскрипционного пузырька возникает небольшой по протяженности и короткоживущий гибрид РНК и ДНК. В составе холофермента РНК-полимеразы, осуществляющего синтез бактериальной РНК, можно выделить два компонента: минимальный фермент — мультимерный комплекс, имеющий субъединичный состав $\alpha_2\beta\beta'$ и осуществляющий элонгацию цепи РНК; и сигма-фактор (σ) — мономер, участвующий в распознавании промотора на стадии инициации.

Минимальный фермент сам по себе обладает сродством к ДНК. В присутствии сигма-фактора аффинность его связывания со случайными последовательностями снижается, а сродство к промоторам, напротив, увеличивается. Скорость, с которой РНК-полимераза находит промоторы, слишком велика, чтобы быть результатом диффузии и случайных контактов с ДНК; вероятнее всего, данный процесс реализуется в виде направленного перебора последовательностей ДНК.

Бактериальные промоторы содержат две короткие консервативные последовательности, центры которых расположены в позициях −35 и −10 (относительно стартовой точки транскрипции). Для большинства промоторов последовательности −35 и −10 очень близки к среднестатистическим консенсусным последовательностям, характеризующим эти регионы. Расстояние, разделяющее консенсусные последовательности, варьирует от 16 до 18 пар оснований. Первоначально РНК-полимераза связывает последовательность −35, а затем взаимодействует с участком −10. В этот момент фермент закрывает от внешних воздействий порядка 77 оснований ДНК. Первичный транскрипционный комплекс переходит из закрытого двойного комплекса в открытый двойной путем расплетания двухцепочечной структуры ДНК на участке длиной примерно 12 пар оснований (расположенном между стартовой точкой транскрипции и участком −10). Важное значение для реакции плавления дуплекса ДНК имеет состав −10 последовательности, богатой основаниями А и Т.

Двойной комплекс превращается в тройной после начала синтеза предшественника рибонуклеотидной цепи. Обычно инициации предшествует ряд абортивных инициаций, в ходе которых РНК-полимераза синтезирует и высвобождает серию очень коротких рибонуклеотидных цепочек, не сдвигаясь при этом с промотора. В конце стадии инициации полимераза меняет свою структуру, теперь минимальный фермент одновременно закрывает около 50 пар оснований. Сигма-фактор при этом высвобождается (в 30% случаев) либо меняет характер связи с минимальным ферментом. После этого минимальный фермент начинает двигаться вдоль ДНК, синтезируя новую цепь РНК. Участок локального расплетания двухцепочечной структуры ДНК движется вместе с полимеразой. К тому моменту, когда растущая цепь достигает длины в 15–20 нуклеотидов, фермент экранирует от внешних воздействий от 30 до 40 пар оснований; и в такой конформации он достигает конца транскрипционной единицы.

Сила промотора характеризует частоту, с которой РНК-полимераза инициирует транскрипцию на данном промоторе; этот параметр напрямую зависит от сходства последовательностей −35 и −10 с идеальной консенсусной последовательностью; влияние оказывают и участки, расположенные по ходу транскрипции сразу за стартовой точкой. Отрицательная сверхспирализация повышает эффективность работы некоторых промоторов. Во время транскрипции образуются положительные сверхвитки ДНК перед РНК-полимеразой и отрицательные сверхвитки позади фермента.

Альтернативные сигма-факторы обеспечивают распознавание минимальным ферментом промоторов с различными консенсусными последовательностями. У *E. coli* эти сигма-факторы активируются под действием неблагоприятных факторов окружающей среды (таких как тепловой шок или азотное голодание). У *B. subtilis* имеется один основной сигма-фактор, обладающий той же специфичностью, что и сигма-фактор *E. coli*, и множество дополнительных сигма-факторов. Другая серия сигма-факторов активируется в процессе споруляции, который регулируется двумя каскадными сменами сигма-факторов в материнской клетке и в предспоре. Механизм регуляции транскрипции посредством смены сигма-факторов реализован также у фага SPO1.

Геометрия комплекса РНК-полимераза–промотор одинакова для холоферментов, содержащих в своем составе любой сигма-фактор (исключение составляет σ^{54}). Каждый сигма-фактор способствует инициации транскрипцию только на промоторах с определенными консенсусными последовательностями −35 и −10. Непосредственный контакт сигма-фактора и ДНК был впервые продемонстрирован для фактора σ^{70} *E. coli.* Этот белок характеризуется наличием N-концевого автоингибирующего пептида, который препятствует ассоциации ДНК-связывающего региона свободного сигма-фактора с последовательностью ДНК. Автоингибирующее дейс-

твие устраняется, когда холофермент формирует открытый комплекс.

Бактериальная РНК-полимераза терминирует транскрипцию в сайтах двух типов. ρ-независимые терминаторы содержат в своем составе шпильку, богатую основаниями G и C, за которой расположен U-богатый регион. Такие терминаторы напрямую распознаются минимальным ферментом *in vitro*. ρ-Зависимые терминаторы для осуществления терминации как *in vivo*, так и *in vitro* требуют участия ρ-фактора; Фактор ρ связывается с так называемыми *rut*-сайтами, содержащими в своем составе много остатков C и мало остатков G, и расположенными перед сайтом терминации. ρ-фактор представляет собой обладающий АТФ-зависимой хеликазной активностью гексамер, перемещающийся вдоль РНК до тех пор, пока не достигнет гетеродуплекса РНК-ДНК. Здесь он способствует диссоциации новосинтезированной цепи РНК из транскрипционного комплекса. Оба типа терминаторных последовательностей вызывают паузу в работе РНК-полимеразы, что необходимо для того, чтобы дать терминационным событиям время на прохождения.

Также в терминации принимают участие факторы Nus. Белок NusA необходим для функционирования ρ-независимых терминаторов, а в случае ρ-зависимых терминаторов требуется участие дополнительного фактора NusB-S10. Димер NusB-S10 распознает последовательность *boxA* в *nut*-сайтах растущей цепи РНК, после чего присоединяется и NusA.

Явление антитерминации используется некоторыми фагами для переключения экспрессии с одной группы генов на другую. Ген N λ-фага кодирует антитерминирующий белок (pN), позволяющий РНК-полимеразе преодолевать терминаторы, расположенные в конце транскрипционных единиц немедленно ранних генов фага. Другой антитерминирующий белок, pQ, активируется на поздних стадиях фаговой инфекции. pN и pQ взаимодействуют с РНК-полимеразой в тот момент, когда она проходит через особые сайты (*nut* и *qut* соответственно). Эти сайты находятся в разных участках своих транскрипционных единиц. Фактор pN узнает РНК-полимеразу в комплексе с фактором NusA, когда комплекс проходит через последовательность *boxB*. pN связывается с транскрипционным комплексом и блокирует терминацию на ρ-зависимом терминаторе, выступая в роли антагониста NusA.

Литература

11.2 Транскрипция происходит на коротком участке расплетенной ДНК

Обзор

Losick, R. and Chamberlin, M. (1976). RNA Polymerase. *Cold Spring Harbor Symp. Quant. Biol.*

Исследование

Korzheva, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A., and Darst, S. A. (2000). A structural model of transcription elongation. *Science* 289, 619–625.

11.3 Реакция транскрипции происходит в три стадии

Исследование

Rice, G. A., Kane, C. M., and Chamberlin, M. (1991).Footprinting analysis of mammalian RNA polymerase II along its transcript: an alternative view of transcription elongation. *Proc. Natl. Acad. Sci. USA* 88, 4245–4281.

Wang, D. et al. (1995). Discontinuous movements of DNA and RNA in RNA polymerase accompany formation of a paused transcription complex. *Cell* 81, 341–350.

11.4 РНК-полимераза фага T7 является удобной модельной системой

Исследование

Cheetham, G. M., Jeruzalmi, D., and Steitz, T. A. (1999). Structural basis for initiation of transcription from an RNA polymerase-promoter complex. *Nature* 399, 80–83.

Cheetham, G. M. T. and Steitz, T. A. (1999). Structure of a transcribing T7 RNA polymerase initiation complex. *Science* 286, 2305–2309.

Temiakov, D., Mentesana, D., Temiakov, D., Ma, K., Mustaev, A., Borukhov, S., and McAllister, W. T. (2000). The specificity loop of T7 RNA polymerase interacts first with the promoter and then with the elongating transcript, suggesting a mechanism for promoter clearance. *Proc. Natl. Acad. Sci. USA* 97, 14109–14114.

11.5 На основе рентгеноструктурного анализа РНК-полимеразы можно построить модель движения фермента

Обзор

Shilatifard, A., Conaway, R. C., and Conaway, J. W. (2003). The RNA polymerase II elongation complex. *Annu. Rev. Biochem.* 72, 693–715.

Исследование

Cramer, P., Bushnell, D. A., Fu, J., Gnatt, A. L., Maier-Davis, B., Thompson, N. E., Burgess, R. R., Edwards, A. M., David, P. R., and Kornberg, R. D. (2000). Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* 288, 640–649.

Cramer, P., Bushnell, P., and Kornberg, R. D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 Å resolution. *Science* 292, 1863–1876.

Gnatt, A. L, Cramer, P., Fu, J., Bushnell, D. A., and Kornberg, R. D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* 292, 1876–1882.

11.6 Бактериальные РНК-полимеразы состоят из нескольких субъединиц

Обзор

Helmann, J. D. and Chamberlin, M. (1988). Structure and function of bacterial sigma factors. *Annu. Rev. Biochem.* 57, 839–872.

Исследование

Campbell, E. A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A., and Darst, S. A. (2001). Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. *Cell* 104, 901–912.

Korzheva, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A., and Darst, S. A. (2000). A structural model of transcription elongation. *Science* 289, 619–625.

Zhang, G., Campbell, E. A., Zhang, E. A., Minakhin, L., Richter, C., Severinov, K., and Darst, S. A. (1999). Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell* 98, 811–824.

11.7 РНК-полимераза состоит из минимального фермента и сигма-фактора

Исследование

Travers, A. A. and Burgess, R. R. (1969). Cyclic reuse of the RNA polymerase sigma factor. *Nature* 222, 537–540.

11.8 После инициации транскрипции характер взаимодействия минимального фермента с сигма-фактором меняется

Исследование

Bar-Nahum, G. and Nudler, E. (2001). Isolation and characterization of sigma(70)-retaining transcription elongation complexes from *E. coli. Cell* 106, 443–451.
Krummel, B. and Chamberlin, M. J. (1989). RNA chain initiation by E. coli RNA polymerase. Structural transitions of the enzyme in early ternary complexes. *Biochemistry* 28, 7829–7842.
Mukhopadhyay, J., Kapanidis, A. N., Mekler, V., Kortkhonjia, E., Ebright, Y. W., and Ebright, R. H. (2001). Translocation of sigma(70) with RNA polymerase during transcription. Fluorescence resonance energy transfer assay for movement relative to DNA. *Cell* 106, 453–463.

11.9 Остановившаяся РНК-полимераза способна продолжить транскрипцию после «перезагрузки»

Исследование

Kettenberger, H., Armache, K. J., and Cramer, P. (2003). Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell* 114, 347–357.
Opalka, N., Chlenov, M., Chacon, P., Rice, W. J., Wriggers, W., and Darst, S. A. (2003). Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase. *Cell* 114, 335–345.

11.11 Сигма-фактор контролирует связывание РНК-полимеразы с ДНК

Исследование

Bar-Nahum, G. and Nudler, E. (2001). Isolation and characterization of sigma(70)-retaining transcription elongation complexes from *E. coli. Cell* 106, 443–451.
Mukhopadhyay, J., Kapanidis, A. N., Mekler, V., Kortkhonjia, E., Ebright, Y. W., and Ebright, R. H. (2001). Translocation of sigma(70) with RNA polymerase during transcription. Fluorescence resonance energy transfer assay for movement relative to DNA. *Cell* 106, 453–463.

11.12 Узнавание промотора зависит от консенсусных последовательностей

Обзор

McClure, W. R. (1985). Mechanism and control of transcription initiation in prokaryotes. *Annu. Rev. Biochem.* 54, 171–204.

Исследование

Ross, W., Gosink, K. K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K., and Gourse, R. L. (1993). A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* 262, 1407–1413.

11.13 В результате мутаций эффективность работы промотора может быть повышена или понижена

Обзор

McClure, W. R. (1985). Mechanism and control of transcription initiation in prokaryotes. *Annu. Rev. Biochem.* 54, 171–204.

11.14 РНК-полимераза связывает молекулу ДНК с одной стороны

Обзор

Siebenlist, U., Simpson, R. B., and Gilbert, W. (1980). *E. coli* RNA polymerase interacts homologously with two different promoters. *Cell* 20, 269–281.

11.15 Сверхспирализация ДНК влияет на процесс транскрипции

Исследование

Wu, H.-Y. et al. (1988). Transcription generates positively and negatively supercoiled domains in the template. *Cell* 53, 433–440.

11.16 Инициация на данном промоторе зависит от типа сигма-фактора

Обзор

Hengge-Aronis, R. (2002). Signal transduction and regulatory mechanisms involved in control of the sigma(S) (RpoS) subunit of RNA polymerase. *Microbiol. Mol. Biol. Rev.* 66, 373–393.

Исследование

Alba, B. M., Onufryk, C., Lu, C. Z., and Gross, C. A. (2002). DegS and YaeL participate sequentially in the cleavage of RseA to activate the sigma(E)-dependent extracytoplasmic stress response. *Genes Dev.* 16, 2156–2168.
Grossman, A. D., Erickson, J. W., and Gross, C. A. (1984). The htpR gene product of *E. coli* is a sigma factor for heat-shock promoters. *Cell* 38, 383–390.
Kanehara, K., Ito, K., and Akiyama, Y. (2002). YaeL (EcfE) activates the sigma(E) pathway of stress response through a site-2 cleavage of anti-sigma(E), RseA. *Genes Dev.* 16, 2147–2155.
Sakai, J., Duncan, E. A., Rawson, R. B., Hua, X., Brown, M. S., and Goldstein, J. L. (1996). Sterol-regulated release of SREBP-2 from cell membranes requires two sequential cleavages, one within a transmembrane segment. *Cell* 85, 1037–1046.

11.17 Сигма-факторы взаимодействуют непосредственно с ДНК

Исследование

Campbell, E. A., Muzzin, O., Chlenov, M., Sun, J. L., Olson, C. A., Weinman, O., Trester-Zedlitz, M. L., and Darst, S. A. (2002). Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol. Cell* 9, 527–539.
Dombrowski, A. J. et al. (1992). Polypeptides containing highly conserved regions of transcription initiation factor σ70 exhibit specificity of binding to promoter DNA. *Cell* 70, 501–512.
Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A. N., Niu, W., Ebright, Y. W., Levy, R., and Ebright, R. H. (2002). Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell* 108, 599–614.

Vassylyev, D. G, Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M. N., Borukhov, S., Yokoyama S. (2002). Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* 417, 712–719.

11.19 Споруляция у бактерий контролируется сигма-факторами

Обзор

Errington, J. (1993). *B. subtilis* sporulation: regulation of gene expression and control of morphogenesis. *Microbiol. Rev.* 57, 1–33.

Haldenwang, W. G. (1995). The sigma factors of *B. subtilis*. *Microbiol. Rev.* 59, 1–30.

Losick, R. and Stragier, P. (1992). Crisscross regulation of cell-type specific gene expression during development in *B. subtilis*. *Nature* 355, 601–604.

Losick, R. et al. (1986). Genetics of endospore formation in *B. subtilis*. *Annu. Rev. Genet.* 20, 625–669.

Stragier, P and Losick, R. (1996). Molecular genetics of sporulation in *B. subtilis*. *Annu. Rev. Genet.* 30, 297–341.

Исследование

Haldenwang, W. G., Lang, N., and Losick, R. (1981). A sporulation-induced sigma-like regulatory protein from *B. subtilis*. *Cell* 23, 615–624.

Haldenwang, W. G. and Losick, R. (1980). A novel RNA polymerase sigma factor from *B. subtilis*. *Proc. Natl. Acad. Sci. USA* 77, 7000–7004.

11.20 Бактериальная РНК-полимераза завершает транскрипцию в определенных сайтах

Обзор

Adhya, S. and Gottesman, M. (1978). Control of transcription termination. *Annu. Rev. Biochem.* 47, 967–996.

Friedman, D. I., Imperiale, M. J., and Adhya, S. L. (1987). RNA 3′ end formation in the control of gene expression. *Annu. Rev. Genet.* 21, 453–488.

Platt, T. (1986). Transcription termination and the regulation of gene expression. *Annu. Rev. Biochem.* 55, 339–372.

11.21 Два типа терминаторов *E. coli*

Обзор

von Hippel, P. H. (1998). An integrated model of the transcription complex in elongation, termination, and editing. *Science* 281, 660–665.

Исследование

Lee, D. N., Phung, L., Stewart, J., and Landick, R. (1990). Transcription pausing by *E. coli* RNA polymerase is modulated by downstream DNA sequences. *J. Biol. Chem.* 265, 15145–15153.

Lesnik, E. A., Sampath, R., Levene, H. B., Henderson, T. J., McNeil, J. A., and Ecker, D. J. (2001). Prediction of ρ-independent transcriptional terminators in *E. coli*. *Nucleic Acids Res.* 29, 3583–3594.

Reynolds, R., Bermadez-Cruz, R. M., and Chamberlin, M. J. (1992). Parameters affecting transcription termination by *E. coli* RNA polymerase. I. Analysis of 13 ρ-independent terminators. *J. Mol. Biol.* 224, 31–51.

11.22 Как работает ρ-фактор?

Обзор

Das, A. (1993). Control of transcription termination by RNA-binding proteins. *Annu. Rev. Biochem.* 62, 893–930.

Richardson, J. P. (1996). Structural organization of transcription termination factor P. *J. Biol. Chem.* 271, 1251–1254.

von Hippel, P. H. (1998). An integrated model of the transcription complex in elongation, termination, and editing. *Science* 281, 660–665.

Исследование

Brennan, C. A., Dombroski, A. J., and Platt, T. (1987). Transcription termination factor ρ is an RNA-DNA helicase. *Cell* 48, 945–952.

Geiselmann, J., Wang, Y., Seifried, S. E., and von Hippel, P. H. (1993). A physical model for the translocation and helicase activities of *E. coli* transcription termination protein P. *Proc. Natl. Acad. Sci. USA* 90, 7754–7758.

Roberts, J. W. (1969). Termination factor for RNA synthesis. *Nature* 224, 1168–1174.

Skordalakes, E. and Berger, J. M. (2003). Structure of the P transcription terminator: mechanism of mRNA recognition and helicase loading. *Cell* 114, 135–146.

11.25 Факторы терминации и антитерминации взаимодействуют с РНК-полимеразой

Обзор

Greenblatt, J., Nodwell, J. R., and Mason, S. W. (1993). Transcriptional antitermination. *Nature* 364, 401–406.

Исследование

Legault, P., Li, J., Mogridge, J., Kay, L. E., and Greenblatt, J. (1998). NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell* 93, 289–299.

Mah, T. F., Kuznedelov, K., Mushegian, A., Severinov, K., and Greenblatt, J. (2000). The alpha subunit of *E. coli* RNA polymerase activates RNA binding by NusA. *Genes Dev.* 14, 2664–2675.

Mogridge, J., Mah, J., and Greenblatt, J. (1995). A protein-RNA interaction network facilitates the template-independent cooperative assembly on RNA polymerase of a stable antitermination complex containing the lambda N protein. *Genes Dev.* 9, 2831–2845.

Olson, E. R., Flamm, E. L., and Friedman, D. I. (1982). Analysis of nutR: a region of phage lambda required for antitermination of transcription. *Cell* 31, 61–70.

Оперон

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

- Петля укладывается в малую бороздку ДНК, как короткая, свернутая α-спираль.
- Индуктор-связывающий сайт и регионы, ответственные за мультимеризацию, расположены в ядре.

12.11 Репрессорный белок — это тетрамер, состоящий из двух димеров

- Мономеры, составляющие димер, формируют контакты между ядерными доменами 1 и 2 и, возможно, между олигомеризационными спиралями.
- Димеры формируют тетрамер посредством контактов между олигомеризационными спиралями.

12.12 Конформация ДНК-связывающего домена регулируется с помощью аллостерических взаимодействий

- ДНК-связывающий домен каждого образующего димер мономера при связывании с ДНК укладывается в большую бороздку ДНК.
- Спираль в шарнире вкладывается в малую бороздку ДНК.
- Активному репрессору свойственна конформация, в которой два ДНК-связывающих домена одного димера могут взаимодействовать с двумя последовательными витками двойной спирали ДНК.
- Связывание индуктора с репрессором приводит к нарушению спиральной структуры петли и изменению конформации ДНК-связывающих сайтов таким образом, что они, находясь в неправильной ориентации, становятся не способны одновременно формировать контакты.

12.13 Мутантные фенотипы коррелируют со структурой отдельных доменов репрессора

- Определенные типы мутаций свойственны отдельным доменам субъединиц репрессора.

12.14 Репрессорный белок связывается с оператором

- Репрессорный белок связывается с двухцепочечной ДНК-последовательностью оператора.
- Оператор представляет собой палиндром длиной 26 пн.
- Каждый инвертированный повтор оператора связывается с ДНК-связывающим сайтом субъединицы репрессора.

12.15 Связывание репрессора с индуктором приводит к его отделению от ДНК

- Связывание репрессора с индуктором приводит к таким изменениям его конформации, в результате которых аффинность связывания с ДНК падает и репрессор отделяется от оператора.

12.16 Репрессор способен связывать три оператора и взаимодействовать с РНК полимеразой

- Каждый димер тетрамерного репрессора может связывать оператор, следовательно, тетрамер может связывать два оператора одновременно.
- Реализация полной репрессии требует, чтобы репрессор провзаимодействовал с дополнительным оператором, расположенным выше или ниже основного оператора, перекрывающегося с промотором гена *lacZ*.
- Связывание репрессора с оператором способствует связыванию РНК полимеразы с промотором, однако препятствует транскрипции.

12.17 Репрессор всегда связан с ДНК

- Белки, обладающие высокой аффинностью к определенным последовательностям ДНК, в отношении всех остальных последовательностей проявляют низкую аффинность.
- Каждая пара оснований бактериального генома может стать началом низкоаффинного сайта связывания для репрессора.
- Наличие большого количества низкоаффинных сайтов связывания гарантирует, что все молекулы репрессорного белка будут ассоциированы с ДНК.
- Репрессор находит оператор, перемещаясь к нему по ДНК от случайного сайта связывания, а не диффундируя из раствора.

12.18 Оператор конкурирует со случайными последовательностями за связывание с репрессором

- В отсутствие индуктора аффинность репрессора к оператору превышает его сродство к случайной последовательности в 10^7 раз.
- Количество молекул репрессора на клетку (около 10) — достаточно для того, чтобы обеспечить связывание оператора в течение 96% времени.
- При добавлении индуктора разница между аффинностью репрессора к оператору и к неспецифическим сайтам сокращается с 10^7 до 10^4. При этом связывается только 3% операторов.
- Присоединение индуктора приводит к смещению репрессора с последовательности оператора на случайную последовательность путем прямой замены одного сайта на другой.
- Эти параметры могут быть изменены путем повышения или понижения эффективной концентрации ДНК *in vivo*.

12.19 Репрессия может происходить сразу в нескольких локусах

- Репрессор может действовать на все участки ДНК, последовательность которых совпадает с последовательностью оператора.

12.20 Цикло-АМФ — эффектор, активирующий БРЦ и опосредующий его действие на многие опероны

- БРЦ — активаторный белок, необходимый для инициации транскрипции на зависимых от него промоторах.
- Для активации димера БРЦ необходима одна молекула циклического АМФ.

12.21 БРЦ функционирует на основе различных механизмов, действуя при этом на разные опероны

- Связываясь с БРЦ, ДНК изгибается на 90°.
- Сайты связывания с БРЦ в разных оперонах-мишенях характеризуются различной локализацией относительно промотора.
- БРЦ взаимодействует с РНК-полимеразой, однако детали этого процесса зависят от положения сайта связывания БРЦ относительно промотора.

12.22 Трансляция может регулироваться

- Белок-репрессор может регулировать трасляцию, препятствуя связыванию рибосомы с инициирующим кодоном.
- С помощью изменений в структуре полицистронной мРНК, возникающих в процессе трансляции, можно контролировать доступность инициирующих кодонов.

12.23 Трансляция рибосомальных белков контролируется аутогенно

- Трансляция рибосомальных белков контролируется одним из продуктов собственного оперона, связывающимся с полицистронной мРНК.

12.24 Синтез белка p32 фага T4 контролируется аутогенно

- p32 связывается со своей собственной мРНК, блокируя инициацию трансляции.

12.25 Аутогенный контроль и сборка макромолекул

- Предшественник белка микротрубочек, свободный тубулин, ингибирует трансляцию тубулиновой мРНК.

12.26 Резюме

12.1 Введение

Контроль уровня экспрессии генов может осуществляться на любом из нескольких этапов реализации генетической информации — транскрипции, процессинга мРЙК и трансляции:

- Регуляция транскрипции осуществляется, как правило, на стадии инициации. Элонгация не задействована в механизмах контроля. Регуляторное влияние стадии терминации проявляется в тот момент, когда РНК-полимераза принимает решение, останавливать транскрипцию на терминаторе или продолжать ее за его пределами.
- В эукариотических клетках процессинг РНК регулируется на стадиях модификации, сплайсинга, транспорта или стабилизации транскриптов. В бактериальных клетках мРНК способна вступать в процесс трансляции сразу же после того, как будет синтезирована (или даже во время синтеза), в этом случае стадия процессинга регуляторным этапом быть не может.
- Трансляция (как и транскрипция) регулируется, как правило, на этапах инициации и терминации. Общие механизмы контроля синтеза белков аналогичны механизмам контроля транскрипции: регуляция стадии инициации и в том, и в другом случае происходит по одной схеме.

Основная концепция регуляции транскрипции у бактерий была сформулирована на основе классической модели оперона, предложенной Жакобом и Моно в 1961 году. Ученые выделили в ДНК два типа последовательностей: последовательности, кодирующие ***транс*-активные** продукты, и ***цис*-активные** последовательности, действующие исключительно на смежные с ними участки ДНК. Активность генов регулируется посредством специфичных взаимодействий *транс*-активных продуктов (обычно — белков) с *цис*-активными последовательностями (обычно — определенных сайтов на ДНК). В более общих терминах:

- Ген — последовательность ДНК, кодирующая способный к диффузии продукт. Этот продукт может быть белком (в большинстве случаев) или РНК (в случае генов, кодирующих тРНК и рРНК). *Основной чертой транс-активного продукта является возможность диффундировать и проявлять свой эффект вдали от места синтеза.* Любой генный продукт, обладающий этим свойством, считается *транс*-активным.
- *Цис*-активной является любая последовательность ДНК, для реализации функций которой не требуется превращения в другую форму (белок, РНК), действующая исключительно *in situ* и оказывающая свое влияние только на смежные с ней участки ДНК. (В некоторых случаях это может быть не только ДНК, но и РНК.)

Для того чтобы различать компоненты регуляторной сети и гены, которые эти компоненты регулируют, мы будем пользоваться терминами — структурные и регуляторные гены. **Структурным геном** считают ген, кодирующий белок (или РНК), обеспечивающий разнообразие структуры и функций внутри клетки, включая структурные белки, ферменты с каталитической активностью и пр. **Регуляторным геном** считают ген, кодирующий белок (или РНК), вовлеченный в регуляцию экспрессии других генов.

Простейшая форма регуляторной модели приведена на РИС. 12.1: *регуляторный ген кодирует белок, контролирующий транскрипцию путем взаимодействия с определенным сайтом или сайтами на ДНК.* Такое взаимодействие может оказывать на ген-мишень как активирующее (повышение уровня экспрессии), так и ингибирующее (снижение уровня экспрессии) влияние. Как правило, участки связывания с регуляторными белками расположены на ДНК непосредственно перед геном-мишенью.

Последовательности ДНК, обозначающие начало и конец транскрипционной единицы (промотор и терминатор), являются примерами *цис*-активных сайтов. *Промотор способен инициировать транскрипцию только одного гена или генов, физически связанных с ним и расположенных в одной и той же молекуле ДНК.* Похожим образом терминатор способен прекращать транскрипцию только той РНК-полимеразы, которая транскрибировала расположенный перед ним ген(ы). Промотор и терминатор представляют собой *цис*-активные элементы, которые распознаются *транс*-активными молекулами к примеру РНК-полимеразой (кроме полимеразы, в процесс могут вовлекаться и другие регуляторные факторы).

Дополнительные *цис*-активные сайты часто примыкают к промоторной последовательности или перемежаются с ней. В бактериальных клетках рядом с промотором (а значит, в непосредственной близости от стартовой точки транскрипции) может быть расположен один или несколько таких сайтов. В эукариотических клетках дополнительных регуляторных участков гораздо больше, и располагаться они могут на значительном удалении друг от друга.

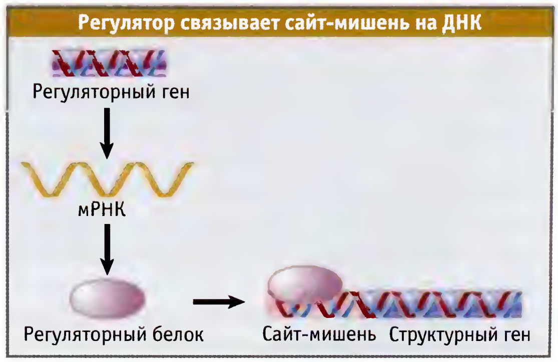


РИС. 12.1 Регуляторный ген кодирует белок, взаимодействующий с определенным участком ДНК

12.2 Негативная и позитивная регуляция экспрессии генов

Ключевые понятия:

- При негативной регуляции белок-репрессор связывается с оператором, блокируя экспрессию гена.
- При позитивной регуляции транскрипционный фактор связывается с промотором, способствуя инициации транскрипции гена РНК-полимеразой.

Классическим примером регуляции у бактерий является *негативная* регуляция с помощью **репрессора**-белка, блокирующего экспрессию гена. На РИС. 12.2 изображена типичная ситуация, при которой экспрессия гена обеспечивается узнаванием РНК-полимеразой его промотора. Рядом с промотором расположен другой *цис*-активный сайт, названный **оператором** и являющийся мишенью для репрессорного белка. Когда репрессор связывается с оператором, РНК-полимераза не может инициировать транскрипцию, и *экспрессия гена выключается.*

Альтернативой негативной регуляции является *позитивная* регуляция. У бактерий такой способ контроля используется (возможно) с той же частотой, что и негативная регуляция, у эукариотических клеток позитивная регуляция встречается чаще. Для инициации транскрипции на промоторе РНК-полимераза требует участия **транскрипционного фактора.** На РИС. 12.3 изображена ситуация, при которой эукариотический ген не активен: РНК-полимераза не может самостоятельно инициировать транскрипцию на промоторе. Несколько *транс*-активных факторов связываются с сайтами-мишенями, расположенными поблизости от промотора. *Взаимодействие с некоторыми из этих факторов способствует инициации транскрипции РНК-полимеразой.*

Обобщив все эти данные, можно сказать, что регуляторный белок — это *транс*-активный фактор, который узнает *цис*-элементы, расположенные (обычно) перед регулируемым геном. Следствием этого взаимодействия становится активация или репрессия гена в зависимости от типа регуляторного белка. Типичным свойством регуляторных белков является их способность распознавать очень короткие последовательности ДНК, как правило, длиной менее 10 пн, несмотря на то что в действительности белок связывается с гораздо большим по размеру участком. Примером может служить бактериальный промотор: РНК-полимераза при инициации экранирует на ДНК > 70 пн, однако необходимым условием узнавания ферментом промотора являются две гексамерные последовательности в положениях −35 и −10.

Существенное отличие в организации генов прокариот и эукариот заключается в том, что бактериальные гены организованы в кластеры, в то время как эукариотические гены функционируют индивидуально. Организация структурных генов в кластеры делает возможным их координированный контроль с помощью одного промотора, в результате чего вся группа генов транскрибируется или не транскрибируется. В этой главе мы обсуждаем механизмы такого контроля и их реализацию у бактерий. Средства координированного контроля нескольких отдельно расположенных эукариотических генов обсуждаются в гл. 25 «Активация транскрипции».

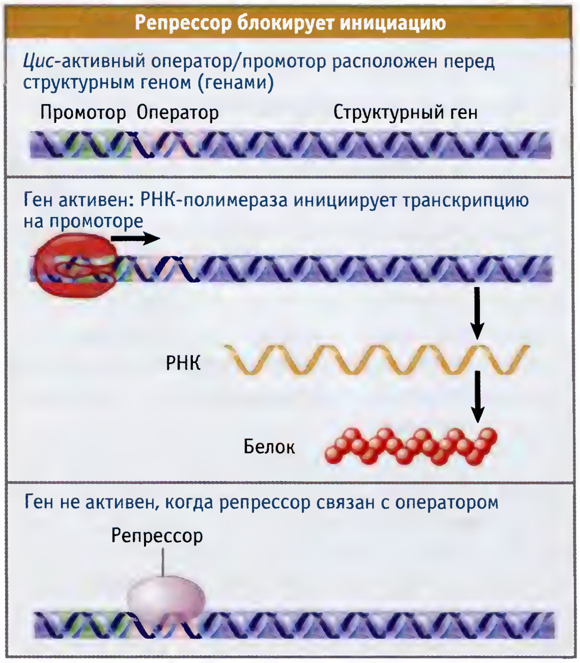


РИС. 12.2 При негативной регуляции *транс*-активный репрессор связывается *цис*-активным оператором и выключает транскрипцию

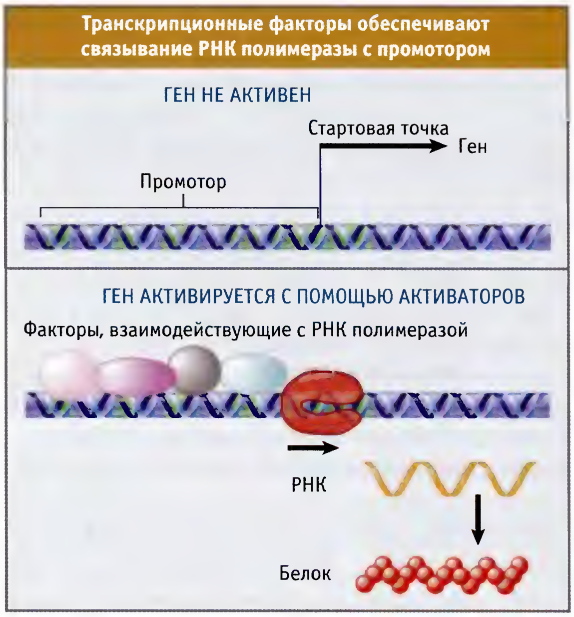


РИС. 12.3 При позитивной регуляции *транс*-активные факторы, связываясь с *цис*-активными сайтами, способствуют инициации транскрипции на промоторе

12.3 Экспрессия генов, организованных в кластеры, регулируется координированно

Ключевые понятия:

- Гены, кодирующие белки со смежными активностями, могут располагаться рядом и контролироваться как одна регуляторная единица, транскрибируемая в одну молекулу полицистронной мРНК.

Структурные гены бактерий, как правило, организованы в кластеры. Обычно кластер включает гены, кодирующие белки со смежными активностями. Такое объединение характерно, например, для генов, кодирующих ферменты одного метаболического пути. Наряду с ферментами, действительно участвующими в определенном процессе биосинтеза, в координированно контролируемую единицу могут быть включены и другие элементы, например белок, ответственный за транспорт малых субстратных молекул в клетку.

Примером такой регуляторной единицы может служить кластер *lacZYA*, состоящий их трех структурных генов *lac*. На РИС. 12.4 представлена организация генов, входящих в этот кластер, ассоциированные с ними *цис*-активные регуляторные элементы и *транс*-активный регуляторный ген. *Ключевым свойством кластера является транскрипция всех составляющих его генов в одну молекулу полицистронной мРНК. Транскрипция всех генов кластера инициируется с одного промотора.*

Ферменты, кодируемые генами *lac*, позволяют клеткам метаболизировать β-галактозиды, к примеру лактозу. Роли трех структурных генов следующие:

- Ген *lacZ* кодирует фермент β-галактозидазу, активная форма которого представляет собой тетрамер с молекулярной массой ~ 500 кДа. Функция фермента заключается в расщеплении молекулы β-галактозида на составляющие ее сахара. К примеру, лактоза расщепляется на глюкозу и галактозу (которые используются в дальнейшем метаболизме). Кроме того, в ходе работы фермента образуется важный промежуточный продукт — β-1,6-аллолактоза, который также участвует в регуляции.
- Ген *lacY* кодирует β-галактозидпермеазу, белок с молекулярной массой 30 кДа, связывающийся с мембраной и являющийся компонентом транспортной системы. Он транспортирует молекулы β-галактозидов в клетку.
- Ген *lacA* кодирует β-галактозидтрансацетилазу, фермент, катализирующий перенос ацетильной группы от ацетил-СоА к β-галактозидам.

Мутации в генах *lacZ* и *lacY* могут приводить к *lac*⁻ генотипу, при котором клетки не могут использовать лактозу. (Обозначение генотипа как «*lac*» без спецификатора означает потерю функции.) Мутации в гене *lacZ* ведут к утрате ферментативной активности, что препятствует сбраживанию лактозы. Мутанты *lacY* не способны использовать лактозу среды. (У клеток *lacA* не обнаружено какого-либо дефекта, и функция этого гена остается невыясненной. Возможно, что реакция ацетилирования дает бактериям преимущество при росте в присутствии определенных неметаболизируемых аналогов β-галактозидов, поскольку эта модификация ведет к их детоксикации и выведению из клетки.)

Система, включающая в себя структурные гены и элементы, контролирующие их экспрессию, формирует полную регуляторную единицу, названную **опероном**. Активность оперона контролируется регуляторным геном (генами), белковый продукт которого взаимодействует с *цис*-активным сайтом-мишенью на ДНК.

12.4 Гены *lac* контролируются репрессором

Ключевые понятия:

- Транскрипция кластера генов *lacZYA* контролируется репрессорным белком, который связывается с оператором — участком ДНК, перекрывающимся с промотором в области стартовой точки транскрипции всего кластера.
- Белок-репрессор — тетрамер, состоящий из четырех одинаковых субъединиц, кодируемых геном *lacI*.

Отличить структурные гены от регуляторных можно по эффекту мутаций. Мутация в структурном гене ведет к отсутствию в клетке определенного белка, кодируемого

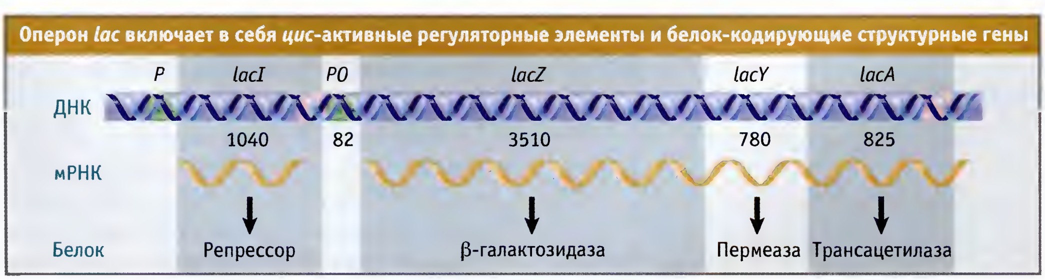


РИС. 12.4 На оперон *lac* приходится ~ 6000 пн ДНК. В левой части расположен ген *lacI*, имеющий свои собственные промотор и терминатор. Конец кодирующей области гена *lacI* непосредственно прилегает к *lacP*-промотору. Оператору *O* соответствуют первые 26 нуклеотидов транскрипционной единицы. Ген *lacZ*, отличающийся большой протяженностью, начинается на 39-м основании, за ним следуют гены *lacZ*, *lacY* и терминатор

этим геном. Мутация же в регуляторном гене влияет на экспрессию всех структурных генов, которые он контролирует. Последствия таких мутаций зависят от типа регуляции.

Транскрипция группы генов *lacZYA* контролируется регуляторным белком, кодируемым геном *lacI*. Ген *lacI* примыкает к структурным генам, но образует независимую транскрипционную единицу со своими собственными промотором и терминатором. В принципе регуляторному гену *lacI* нет необходимости располагаться рядом со структурными генами, поскольку он кодирует диффундирующий продукт. Он способен функционировать с равной эффективностью, располагаясь на любом удалении от генов-мишеней, даже в случае локализации на другой молекуле ДНК (классический тест для *транс*-активного регулятора).

Экспрессия генов *lac* контролируется по принципам **негативной регуляции**: *они экспрессируются до тех пор, пока транскрипция не будет блокирована регуляторным белком.* В результате мутации, инактивирующей регулятор, структурные гены лактозного оперона остаются в активном состоянии. Поскольку функция продукта регуляторного гена *lacI* заключается в блокировании экспрессии структурных генов, он был назван *Lac-репрессором.*

Белок-репрессор представляет собой тетрамер, состоящий из идентичных субъединиц с молекулярной массой 38 кДа каждая. На одну клетку дикого типа приходится ~ 10 таких тетрамеров. Работа регуляторного гена не зависит от наличия в среде субстрата — лактозы. Он транскрибируется в моноцистронную мРНК со скоростью, которая определяется степенью сродства его промотора к РНК-полимеразе.

Репрессор связывается с оператором (обозначаемым как O_{lac}) в стартовой точке кластера *lacZYA*. Оператор располагается между промотором (P_{lac}) и структурными генами (*lacZYA*). *Если репрессор связывается с оператором, его присутствие блокирует инициацию транскрипции на промоторе.* На рис. 12.5 представлено схематичное изображение региона, расположенного в районе стартовой точки структурных генов *lac*. Оператор расположен между основаниями −5 и +21, перекрываясь, таким образом, с правым концом промотора. Более детально взаимодействие между репрессорным белком и РНК-полимеразой обсуждается в разд. 12.14 «Репрессорный белок связывается с оператором» и разд. 12.16 «Репрессор связывает три оператора и взаимодействует с РНК-полимеразой».

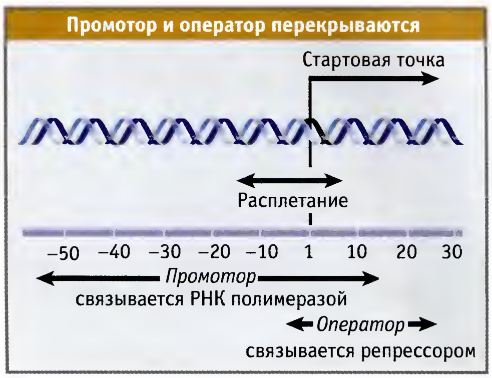


Рис. 12.5 Репрессор и РНК-полимераза связываются с сайтами оперона *lac*, перекрывающимися в районе стартовой точки транскрипции

12.5 Индукция *lac* оперона

Ключевые понятия:

- Молекулы, индуцирующие оперон, либо являются субстратом для ферментов, кодируемых этим опероном, либо обладают какими-то общими с потенциальным субстратом свойствами.
- β-галактозид — субстрат для ферментов, кодируемых *lacZYA*.
- В отсутствие β-галактозида гены *lac*-оперона экспрессируются на очень низком (фоновом) уровне.
- Добавление определенных β-галактозидов индуцирует транскрипцию всех трех генов *lac*-оперона.
- *lac*-мРНК крайне нестабильна, в связи с чем эффект индукции наблюдается непродолжительное время.
- Индукция оперонов метаболических ферментов субстратами этих ферментов и репрессия оперонов биосинтетических ферментов продуктами ферментативной активности происходит по одним и тем же механизмам.

Бактериям необходимо уметь быстро отвечать на изменения условий окружающей среды. Доступность тех или иных питательных веществ может постоянно меняться, и выживаемость в этом случае будет зависеть от способности переключать метаболизм с одного субстрата на другой. Экономичность тоже имеет большое значение: бактерия, которая в ответ на изменения условий окружающей среды использует пути, требующие больших энергетических затрат, может оказаться в невыгодном положении. Бактерия не синтезирует ферментов того или иного пути в отсутствие соответствующего субстрата, но способна в любое время начать синтез при его появлении.

Синтез фермента в ответ на появление в среде соответствующего субстрата получил название **индукции**. Этот тип регуляции широко распространен у бактерий и одноклеточных эукариот (дрожжи). Контролирующий механизм такого типа характерен для организации генов лактозной системы *E. coli*.

При выращивании клеток *E. coli* в отсутствие β-галактозида, они производят очень небольшое число молекул фермента β-галактозидазы, необходимого для метаболизма β-галактозидов (менее 5). При добавлении соответствующего субстрата ферментативная активность проявляется в бактериальных клетках очень быстро. Первые молекулы фермента появляются в течение 2–3 минут, и вскоре их количество достигает > 5000 молекул фермента на одну бактериальную клетку. (При определенных условиях на долю β-галактозидазы может приходится 5–10% от общего растворимого белка клет-

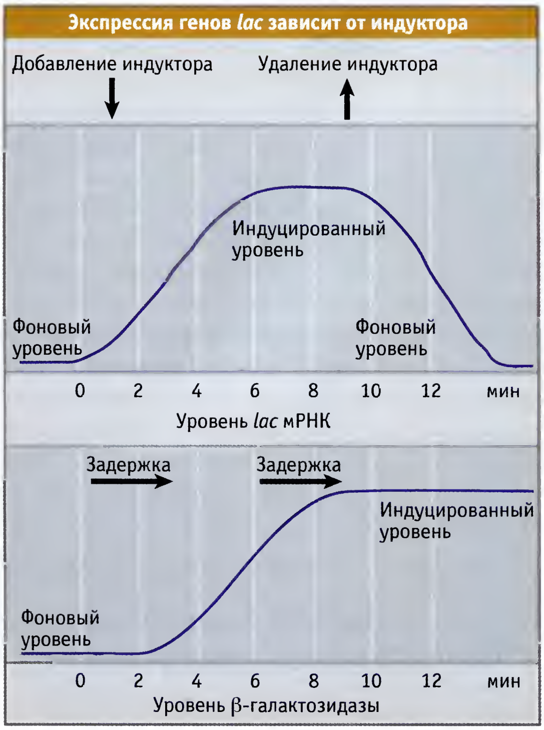


РИС. 12.6 Добавление индуктора приводит к быстрой индукции синтеза *lac* мРНК, после чего следует небольшая задержка, после чего повышается и уровень продукции фермента; сразу после удаления индуктора происходит остановка синтеза

ки.) При удалении субстрата из среды синтез фермента прекращается так же быстро, как был начат.

На РИС. 12.6 приведены основные черты процесса индукции. Как показано в верхней части рисунка, уровень экспрессии генов *lac* быстро меняется в ответ на индукцию. В отсутствие индуктора оперон транскрибируется на очень низком **фоновом уровне**. Добавление индуктора стимулирует транскрипцию; количество *lac*-мРНК быстро поднимается до индуцированного уровня, что отражает баланс между синтезом и деградацией мРНК.

lac-мРНК крайне нестабильна, период ее полураспада составляет ~3 минуты. Это свойство мРНК позволяет быстро снимать эффект индукции. Транскрипция прекращается, как только удаляется индуктор. В очень короткий промежуток времени вся лактозная мРНК деградирует и ее уровень в клетке возвращается к фоновому.

Динамика синтеза белка приведена в нижней части рисунка. В результате трансляции *lac*-мРНК образуется β-галактозидаза (и продукты других лактозных генов). Задержка между появлением *lac*-мРНК и появлением первых молекул полного фермента очень небольшая (между моментом превышения уровня *lac*-мРНК над фоновым и увеличением количества белка проходит ~ 2 минуты). Примерно столько же времени проходит между точками, когда уровни мРНК и белка достигают максимума. При удалении индуктора синтез фермента почти немедленно прекращается (как только деградирует мРНК). Молекулы β-галактозидазы более стабильны, чем молекулы мРНК, поэтому ферментативная активность сохраняется на индуцированном уровне дольше.

Такой тип быстрого ответа на изменения в составе питательной среды наблюдается не только при необходимости использовать новые субстраты. Он используется также для выключения синтеза эндогенных соединений при их внезапном появлении в среде. Например, *E. coli* синтезирует аминокислоту триптофан при участии фермента триптофан-синтазы. Однако, если в среде, на которой растут бактерии, присутствует триптофан, синтез фермента немедленно прекращается. Это явление получило название **репрессии**. Оно позволяет бактериальной клетке избежать траты своих ресурсов на ненужную в данный момент синтетическую активность.

Индукция и репрессия представляют собой разные стороны одного и того же явления. В одном случае бактериальная клетка регулирует свою способность к использованию имеющегося субстрата (лактозы); в другом она регулирует способность к синтезу определенного промежуточного компонента метаболического пути (необходимой аминокислоты). Каждый тип регуляции опосредуется малыми молекулами, которые являются либо субстратом для фермента, либо продуктом ферментативной активности соответственно. Малые молекулы, индуцирующие образование метаболизирующих их ферментов, называют **индукторами**. Тех же, которые предотвращают образование синтезирующих их ферментов, называют **корепрессорами**.

12.6 Малые молекулы индуктора контролируют активность репрессора

Ключевые понятия:

- Индукторы действуют на репрессорный белок, понижая аффинность его связывания с оператором.
- Репрессор имеет два сайта связывания: один для связывания с оператором, другой — с индуктором.
- Репрессор инактивируется в результате аллостерических взаимодействий с индуктором, изменяющих свойства его ДНК-связывающего сайта.

Способность соединений действовать в качестве индуктора или корепрессора высокоспецифична. В процесс могут включаться только субстраты/продукты или близкородственные молекулы. В большинстве случаев *активность самого индуктора не зависит от его взаимодействия с ферментом.* Природным индуктором для системы *lac* служит побочный продукт работы фермента lacZ — β-1,6-аллолактоза, образующийся в результате трансгликозилирования β-1,4-связи в лактозе. Кроме того, аллолактоза является субстратом для lacZ и не сохраняется в клетке. Некоторые молекулы очень похожи на природные индукторы *lac*-оперона, но не могут метаболизироваться ферментом. Ярким приме-

ром может служить изопропилтиогалактопиранозид (ИПТГ), один из нескольких тиогалактозидов с такими свойствами. ИПТГ является очень эффективным индуктором, хотя и не узнается β-галактозидазой.

Молекулы, индуцирующие синтез ферментов, но не метаболизируемые ими, получили название **неметаболизируемых индукторов**. Они оказались крайне полезными благодаря их способности сохраняться в клетке в исходном состоянии. (Истинный индуктор был бы метаболизирован, что препятствовало бы изучению системы.) Существование стабильных индукторов позволило установить важный факт. *Помимо ферментов, система должна включать в себя дополнительные компоненты, способные узнавать соответствующий субстрат; эта способность (узнавать родственные потенциальные субстраты) отличается от таковой у фермента.*

Компонентом системы, отвечающим на индуктор, является репрессорный белок, кодируемый геном *lacI*. Структурные гены *lacZYA* транскрибируются в одну молекулу мРНК, транскрипция начинается с промотора, расположенного непосредственно перед генов *lacZ*. От работы репрессора зависит, произойдет ли инициация транскрипции на промоторе:

- Как показано на РИС. 12.7, в отсутствие индуктора гены не транскрибируются, поскольку репрессор в своей активной форме связан с оператором.
- Как показано на РИС. 12.8, при добавлении индуктора репрессор превращается в форму, которая не способна связываться с оператором, либо в форму, которая, обладая пониженной аффинностью к оператору, отделяется от него. После этого транскрипция инициируется на промоторе и продолжается до терминатора, расположенного за 3′-концом гена *lacA*.

Критическое значение для регуляторной системы имеет способность репрессора блокировать транскрипцию и узнавать малые молекулы индуктора. Репрессор обладает двумя участками связывания, один предназначен для оператора, другой — для индуктора. При связывании индуктора со своим участком конформация репрессора изменяется таким образом, что оказывает влияние на активность сайта связывания с оператором. Такой тип взаимодействия получил название **аллостерического** контроля.

Индукция опосредует **координированную регуляцию**, при которой *все гены экспрессируются (или не экспрессируются) одновременно*. Трансляция мРНК осуществляется последовательно с 5′-конца. Это объясняет, почему при индукции сначала появляется β-галактозидаза, затем β-галактозид-пермеаза и, наконец, β-галактозид-трансацетилаза. Трансляция продуктов трех генов с одной молекулы мРНК объясняет также, почему относительные количества трех ферментов остаются постоянными при варьирующих условиях индукции.

Индукция служит сигналом к экспрессии генов. Индукторы могут отличаться по своей эффективности. Кроме того, на абсолютный уровень транскрипции и трансляции могут влиять и другие факторы, однако соотношение между продуктами трех генов предопределено организацией генов.

Можно заметить некий парадокс в организации работы оперона. Лактозный оперон содержит в своем составе структурный ген (*lacZ*), который кодирует фермент β-галактозидазу, необходимый для метаболизма

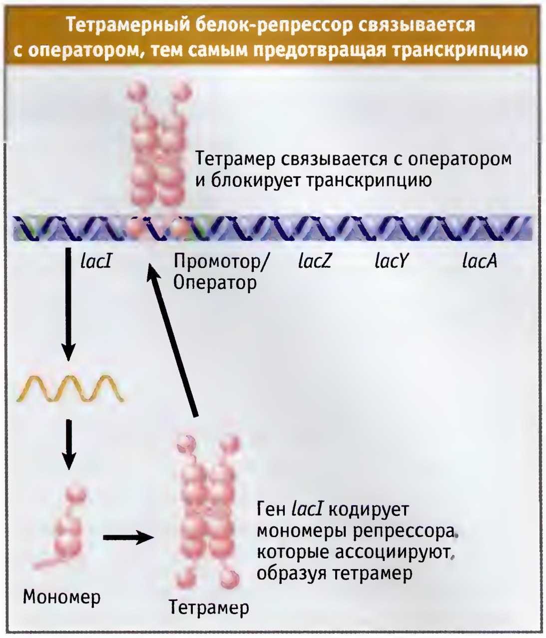


РИС. 12.7 Репрессор, связанный с оператором, поддерживает *lac*-оперон в неактивном состоянии. Как показывает структура кристалла, репрессорный белок образован серией связанных доменов (см. ниже)

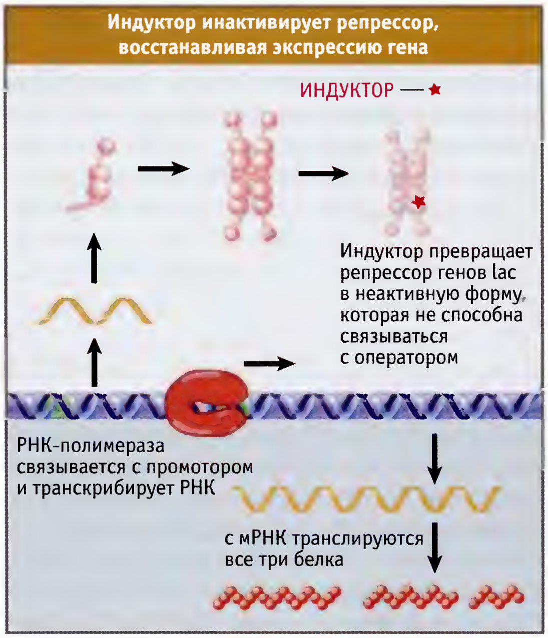


РИС. 12.8 При добавлении индуктора репрессор переходит в неактивную форму, неспособную связываться с оператором. Это позволяет РНК-полимеразе беспрепятственно инициировать транскрипцию

сахара; он также включает в себя ген *lacY*, кодирующий белок, необходимый для транспорта субстрата в клетку. Как же индуктор проникает в клетку и начинает процесс индукции в том случае, если оперон находится в репрессированном состоянии?

Две особенности организации оперона в любой момент обеспечивают минимальное количество белка в клетке, достаточное для того, чтобы запустить процесс. Во-первых, это экспрессия оперона на фоновом уровне: даже в отсутствие индукции он экспрессируется на низком уровне (0,1% от индуцированного уровня). Во-вторых, некоторые индукторы способны проникать в клетку без участия специальной системы транспорта.

12.7 Оператор можно обнаружить с помощью *цис*-активных мутаций

Ключевые понятия:

- Результатом мутаций в операторе является конститутивная экспрессия всех трех структурных генов *lac*-оперона.
- Эффект *цис*-активных мутаций сказывается только на генах, расположенных на смежном с оператором участке ДНК.

Мутации, затрагивающие элементы регуляторной сети, могут блокировать экспрессию оперона либо вызывать неконтролируемую его экспрессию. Опероны, которые в результате мутации теряют способность экспрессироваться, называются **неиндуцируемыми**. Экспрессию генов, осуществляемую независимо от действия регуляторов, называют **конститутивной** экспрессией. Мутантов, обладающих таким свойством, называют конститутивными мутантами.

Компоненты регуляторной сети могут быть обнаружены с помощью мутаций, *влияющих на экспрессию всех структурных генов, располагаясь при этом за их пределами*. Такие мутации можно разделить на два класса: промотор и оператор как мишени для регуляторных белков (РНК-полимераза и репрессор, соответственно). Мутации, блокирующие выработку *транс*-активного продукта, позволили идентифицировать локус *lacI* как ген репрессорного белка.

Исходно оператор был обнаружен благодаря конститутивным мутациям (обозначаемым O^c), свидетельствующим, что эта область, *поврежденная мутациями, не чувствует диффундирующего продукта*.

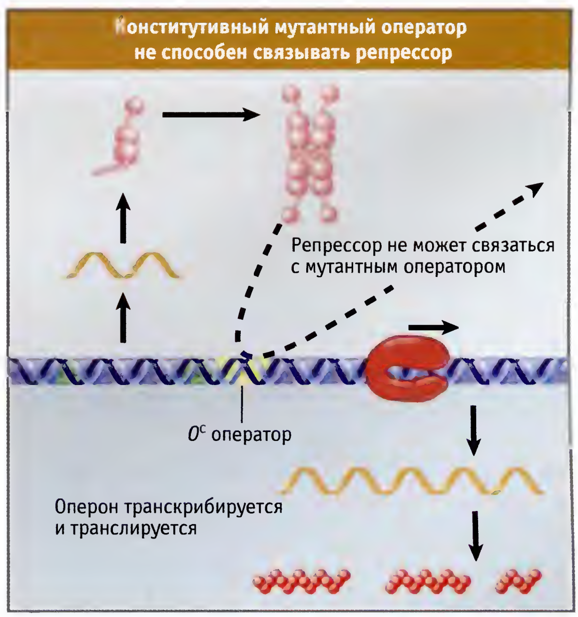
Структурные гены, смежные с местом мутации O^c, экспрессировались конститутивно, что объясняется мутационными изменениями оператора, предотвращающими связывание с ним репрессора. В результате репрессор оказывается не способен препятствовать инициации транскрипции РНК-полимеразой. Оперон экспрессируется конститутивно, как это показано на РИС. 12.9.

Оператор способен контролировать только расположенные рядом с ним гены lac. Если в бактериальную клетку ввести второй *lac*-оперон (на независимой молекуле ДНК), который будет содержать свой собственный оператор, два оператора не будут влиять друг на друга. В том случае, когда оперон имеет оператор дикого типа, он будет репрессироваться при обычных условиях, тогда как второй оператор с мутацией O^c будет работать характерным для него путем.

Эти свойства позволяют определить оператор как типичный *цис*-активный сайт, чьи функции зависят от узнавания его последовательности определенными *транс*-активными факторами. Оператор контролирует смежные с ним гены независимо от присутствия в клетке других аллелей данного сайта. Мутации в таком сайте — к примеру, O^c-мутация — называют ***цис*-доминантными**.

Для мутации, произошедшей в *цис*-активном сайте, невозможно определить принадлежность к группе комплементации. (Способность к комплементации характерна только для генов, экспрессируемых с образованием диффундирующего продукта.) Когда два *цис*-активных сайта расположены близко друг от друга — например, промотор и оператор, — мы не можем классифицировать мутации с помощью комплементационного теста. Мы можем различить их только по влиянию на фенотип.

Цис-доминантность характерна для любого сайта, *физически связанного с контролируемой последовательностью*. Если контролирующий сайт функционирует как часть полицистронной мРНК, мутации в нем будут проявлять точно *такой же тип цис-доминантности*, как если бы он функционировал в ДНК. Важно отметить, что



Конститутивные мутации в операторе блокируют его связывание с репрессором, в результате чего РНК-полимераза получает доступ к промотору. Мутации O^c влияют только на смежные с местом мутации структурные гены и потому относятся к *цис*-активным

контролирующий сайт в рассмотренном примере должен быть физически сцеплен с генами, которые он регулирует. Причем с генетической точки зрения не имеет значения, находятся сайт и гены вместе на ДНК или на РНК.

12.8 Регуляторный ген можно выявить с помощью *транс*-активных мутаций

Ключевые понятия:

- *Транс*-активные мутации гена *lacI* оказывают эффект на экспрессию всех генов кластера *lacZYA* бактерии.
- Мутации, блокирующие функцию гена *lacI*, рецессивны и приводят к конститутивной экспрессии.
- Мутации в ДНК-связывающем сайте репрессорного белка конститутивны, так как приводят к потере репрессором способности связываться с оператором.
- Мутации в индуктор-связывающем сайте репрессорного белка блокируют процесс его инактивации и делают индукцию невозможной.
- *Цис*-активные мутации в промоторе приводят к неиндуцируемости оперона.

Мутации *lacI*⁻ типа, вызванные потерей функции (включая делецию гена), приводят к конститутивной транскрипции. Транскрипция может инициироваться на промоторе в случае, когда репрессор неактивен или отсутствует. Как показано на РИС. 12.10, *lacI*⁻ мутанты экспрессируют все три структурных гена постоянно (конститутивно), *независимо от того, присутствует индуктор или нет*, поскольку репрессор неактивен.

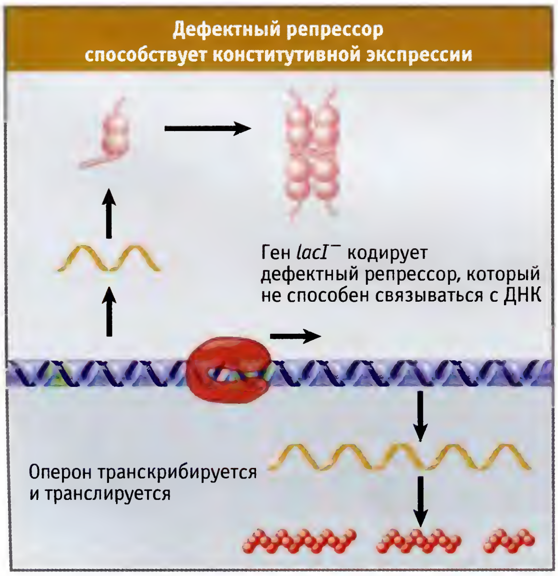


РИС. 12.10 Мутации, инактивирующие ген *lacI*, приводят к тому, что репрессорный белок теряет способность связываться с оператором, а экспрессия оперона приобретает конститутивный характер

С генетической точки зрения, конститутивные мутации могут быть разделены на два класса. O^c-мутанты доминантны, в то время, как *lacI*⁻ — рецессивны. Это означает, что введение гена *lacI*⁺ восстанавливает нормальную регуляцию, на которую не влияет присутствие дефектного *lacI*⁻ гена.

Мутантов, конститутивные мутации в которых приводят к неспособности оперона отвечать на индукцию, также делят на два генетических класса:

- Мутации в промоторе *цис*-активны. Если мутация препятствует связыванию РНК-полимеразы с P_{lac}, оперон становится нефункциональным, поскольку не может транскрибироваться.
- Мутации, блокирующие способность репрессора связываться с индуктором, описываются, как *lacI*ˢ. Они *транс*-активны. Репрессор «зафиксирован» в активной форме, которая узнает оператор и препятствует транскрипции. Добавление индуктора не оказывает должного эффекта, поскольку его связывание с репрессором блокировано, либо связывание не приводит к превращению репрессора в форму с пониженной аффинностью к оператору. Мутантный репрессор связывается со всеми *lac*-операторами в клетке, препятствуя их транскрипции, и не может быть отделен от них, независимо от свойств любого другого белка-репрессора дикого типа, который может присутствовать в клетке, то есть такие мутации доминируют над диким типом.

Эти два типа мутаций в гене *lacI* могут быть использованы для выявления активных центров репрессорного белка. *ДНК-связывающий сайт* репрессора распознает последовательность оператора. Он был определен с помощью конститутивных точечных мутаций, блокирующих связывание репрессора с ДНК, в результате чего тот терял способность останавливать работу РНК-полимеразы. *Индуктор-связывающий сайт репрессора* был идентифицирован с помощью точечных мутаций, вызывающих неиндуцибельность оперона. В результате таких мутаций индутор не мог связываться с репрессором и запускать аллостерические изменения в конформации ДНК-связывающего сайта.

12.9 Мультимерные белки обладают особыми генетическими свойствами

Ключевые понятия:

- Активный репрессор представляет собой тетрамер, состоящий из идентичных субъединиц.
- В случае когда присутствуют и мутантные субъединицы и субъединицы дикого типа, единственная мутантная субъединица *lacI*$^{-d}$ может инактивировать тетрамер, в котором все остальные субъединицы — дикого типа.
- Мутации *lacI*$^{-d}$ затрагивают ДНК-связывающий сайт. Эффект этих мутаций объясняется тем, что для реализации репрессорной активности необходимо, чтобы все ДНК-связывающие сайты в тетрамере были активны.

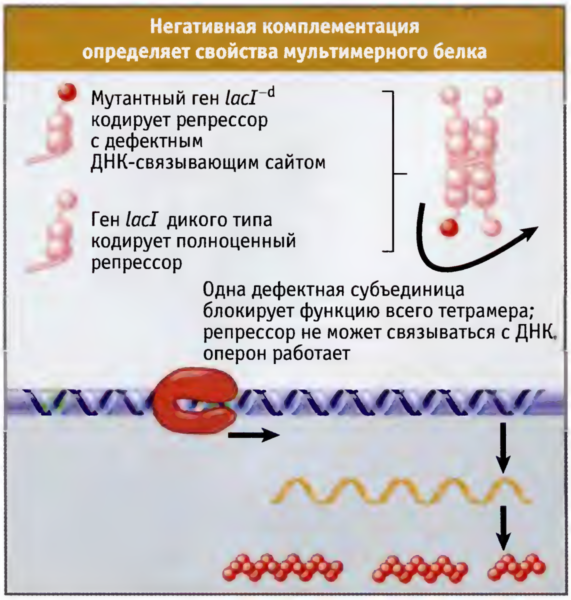


РИС. 12.11 Мутантный аллель $lacI^{-d}$ кодирует субъединицу репрессора с дефектным ДНК-связывающим сайтом (отмечено красным шариком). Когда этот аллель присутствует в клетке одновременно с аллелем дикого типа, мультимерный репрессорный белок собирается случайным образом из мономеров обоих типов. Для того чтобы функция репрессора была блокирована, достаточно, чтобы в его мультимерной структуре присутствовала всего одна субъединица $lacI^{-d}$. Этим объясняется явление доминантно-негативной активности аллеля $lacI^{-d}$

Важной особенностью репрессора является его мультимерная структура. Субъединицы репрессора ассоциируют в клетке случайным образом, формируя активный тетрамерный белок. В случаях, когда в клетке имеется два разных аллеля гена *lacI*, субъединицы двух типов могут соединяться с образованием гетеротетрамера, который по своим свойствам может отличаться от гомотетрамера. Такой тип взаимодействия между субъединицами называют **межаллельной комплементацией**, являющейся характерным свойством мультимерных белков.

В случае некоторых мутантных репрессоров имеет место определенный тип взаимодействия, названный **негативной комплементацией**. Она может наблюдаться при комбинировании генов $lacI^{-d}$ и $lacI^{+}$. Мутация $lacI^{-d}$ приводит к образованию репрессора, который не способен связываться с оператором и относится поэтому к аллелям конститутивного типа. Мутация типа $lacI^{-}$ инактивирует репрессор и является рецессивной по отношению к аллелю дикого типа. Однако индекс *–d* указывает на то, что этот вариант негативного типа проявляет доминантность в случае объединения с аллелем дикого типа. Такие мутации называют **доминантно негативными**.

На РИС. 12.11 представлено объяснение этого феномена. Доминирование обусловлено тем, что аллель $lacI^{-d}$ кодирует дефектную субъединицу, которая не только сама по себе не способна связываться с оператором, но, являясь частью тетрамера, препятствует также связыванию с ним любой полноценной субъединицы. Из этого следует, что для реализации репрессии необходим репрессор-тетрамер, а не отдельный мономер. Действительно, во всех случаях, когда белок имеет доминантно негативную форму, это должно означать, что он функционирует как часть мультимерного комплекса. Использование белков с доминантно негативной активностью стала популярным приемом в генетике эукариот.

12.10 Субъединица репрессора состоит из нескольких доменов

Ключевые понятия:

- В структуре субъединицы репрессорного белка могут быть выделены N-концевой ДНК-связывающий домен, шарнир и ядро.
- ДНК-связывающий домен представлен двумя короткими α-спиралями, которые связывают большую бороздку ДНК.
- Петля укладывается в малую бороздку ДНК как короткая свернутая α-спираль.
- Индуктор-связывающий сайт и регионы, ответственные за мультимеризацию, расположены в ядре.

В структуре мономерной субъединицы репрессора выделяют несколько доменов. На ДНК-связывающий домен приходятся аминокислотные остатки 1-59, этот участок назван длинным N-концевым фрагментом, или **головным участком** (headpieces). Под действием трипсина он может быть отщеплен от другой части мономера, называемой *ядром*. Структура кристалла, приведенная на РИС. 12.12 дает возможность рассмотреть эти регионы более детально.



РИС. 12.12 В структуре мономера Lac-репрессора выделяют несколько независимых доменов. Схема была составлена Hangli Zhan на основании файла 1lbg из базы данных PDB (Protein Data Base) и предоставлена Kathleen S. Matthews, Rice University

N-концевой фрагмент мономера представляет собой две α-спирали, разделенные одним витком. Эта структура является общей для всех ДНК-связывающих мотивов и называется HTH (helix-turn-helix); две α-спирали связываются с ДНК в большой бороздке, где они могут образовывать контакты с определенными основаниями (см. разд. 14.11 «Мотив «спираль-поворот-спираль» необходим репрессору для связывания ДНК»). Данный регион посредством *шарнира* соединен с основной частью белковой молекулы. В ДНК-связанной форме репрессора шарнир формирует короткую α-спираль (как показано на рисунке). Когда же репрессор не связан с ДНК, структура этого региона не упорядочена. HTH-мотив и шарнир вместе образуют ДНК-связывающий домен.

Большая часть ядра приходится на два региона с одинаковыми структурами (ядерные домены 1 и 2). Каждый представлен шестицепочечной β-складчатой структурой, расположенной между двумя α-спиралями. Между двумя этими регионами находится щель, в которую входит индуктор.

С-концевой участок представляет собой α-спираль, состоящую из двух лейциновых гептамерных повторов, создающих олигомеризационный домен. Олигомеризационные спирали четырех мономеров ассоциируют с образованием тетрамерной структуры.

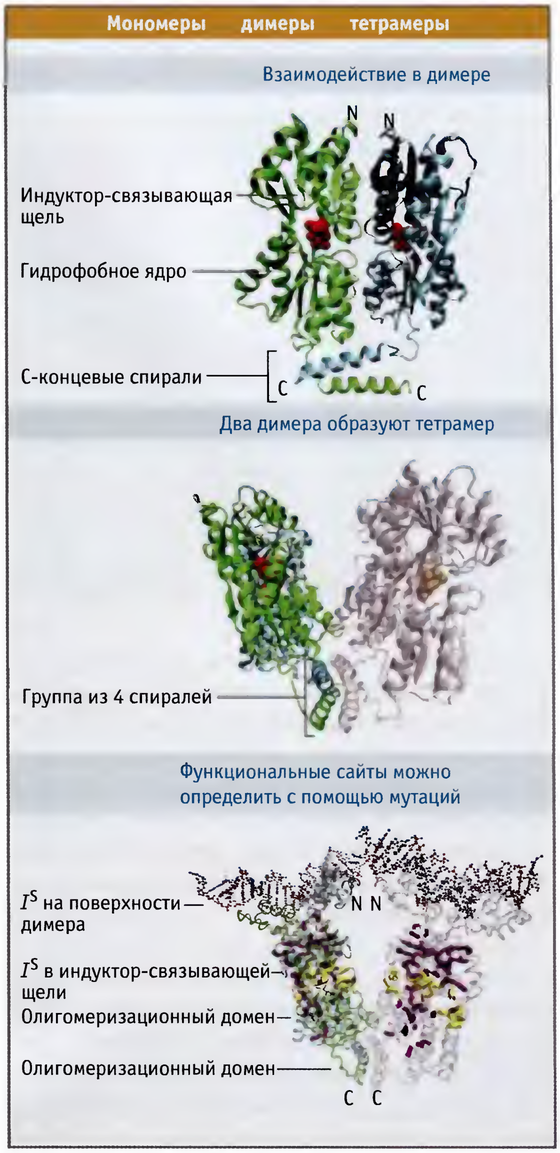


РИС. 12.13 Структура кристалла ядра Lac-репрессора отражает природу взаимодействия между мономерами в тетрамере. Каждый мономер выделен отдельным цветом. Обозначения мутаций: мутации в участках, соответствующих поверхности димера, выделены желтым цветом, в индуктор-связывающем сайте — синим, в олигомеризационном домене — белым и пурпурным. Любезно предоставлено Benjamin Wieder и Ponzy Lu, Univesity of Pennsylvania

12.11 Репрессорный белок — это тетрамер, состоящий из двух димеров

Ключевые понятия:

- **Мономеры, составляющие димер, формируют контакты между ядерными доменами 1 и 2 и, возможно, между олигомеризационными спиралями.**
- **Димеры формируют тетрамер посредством контактов между олигомеризационными спиралями.**

На РИС. 12.13 изображена структура тетрамерного ядра белка-репрессора (построенная на основе структуры, приведенной на рис. 12.12). Она включает в себя два димера. В димере, в свою очередь, различают N-концевые участки мономеров, соприкасающиеся своими поверхностями, щель, в которой связывается индуктор, и гидрофобное ядро (верхняя часть рисунка). С-концевые регионы каждого мономера выпячиваются в виде параллельных спиралей. Головной (длинный N-концевой фрагмент) участок соединяется с N-концевыми регионами в верхней части рисунка. Взаимодействуя, димеры образуют тетрамер (центральная часть рисунка), который удерживает свою структуру благодаря взаимодействиям С-концевых спиральных участков.

Сайты мутаций показаны на структуре тетрамера точками (нижняя часть рисунка). Мутации *lacI^S* делают репрессор нечувствительным к индуктору, а оперон — неиндуцибельным. Они расположены группами: серым цветом обозначены мутации, картируемые в индуктор-связывающем участке, желтым — затрагивающие поверхность димера в области ядерного домена 1. Первая группа мутаций блокирует индуктор-связывающий сайт; вторая нарушает аллостерические взаимодействия между индуктор-связывающим и ДНК-связывающим сайтами репрессора. Мутации *lacI^{-d}*, влияющие на процесс олигомеризации, также расположены двумя группами. Белым цветом показаны мутации в ядерном домене 2. Эти мутации препятствуют формированию димера. Пурпурным цветом показаны мутации в олигомеризационной спи-

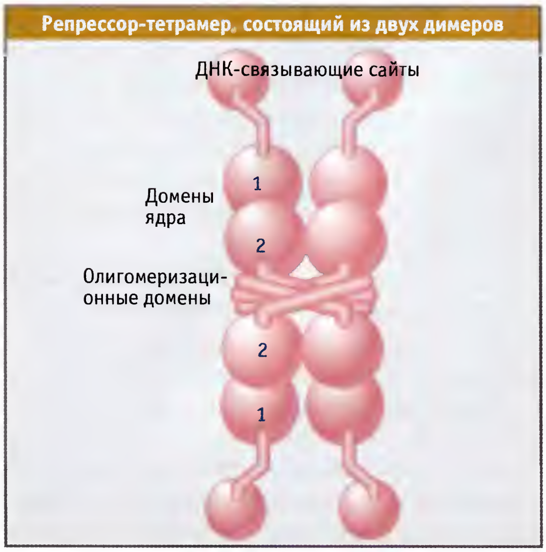


РИС. 12.14 Тетрамерный репрессор состоит из двух димеров. Димеры удерживаются вместе с помощью контактов, образующихся между ядерными доменами 1 и 2, а также между олигомеризационными спиралями. Димеры соприкасаются друг с другом олигомеризационными поверхностями

рали, нарушающие процесс образования тетрамера из димеров.

На основе изложенных выше данных мы составили схему, которая демонстрирует, как мономеры репрессора организованы в тетрамерный белок. Схема приведена на РИС. 12.14. Два мономера формируют димер, соединяясь ядерными доменами 2 и олигомеризационными спиралями. Димер имеет два ДНК-связывающих участка с одной стороны комплекса и олигомеризационными спиралями с другой стороны. После этого два димера организуются в тетрамер, взаимодействуя своими олигомеризационными поверхностями.

12.12 Конформация ДНК-связывающего домена регулируется с помощью аллостерических взаимодействий

Ключевые понятия:

- ДНК-связывающий домен каждого образующего димер мономера при связывании с ДНК укладывается в большую бороздку ДНК.
- Спираль в шарнире вкладывается в малую бороздку ДНК.
- Активному репрессору свойственна конформация, в которой два ДНК-связывающих домена одного димера могут взаимодействовать с двумя последовательными витками двойной спирали ДНК.
- Связывание индуктора с репрессором приводит к нарушению спиральной структуры петли и изменению конформации ДНК-связывающих сайтов таким образом, что они, находясь в неправильной ориентации, становятся неспособны одновременно формировать контакты.

В ранних работах, посвященных изучению взаимодействия репрессоров с ДНК, постулируется модель, согласно которой головной фрагмент молекулы-репрессора относительно независим от ядра. Он может связываться с последовательностью оператора на ДНК, образуя те же самые контакты, что и интактный белок, хотя аффиность его связывания с ДНК намного меньше, в сравнении с неповрежденным белком. Эта разница объясняется тем, что димерная форма интактного репрессора позволяет сразу двум N-концевым фрагментам образовывать контакты с оператором. На РИС. 12.15 показано, как два ДНК-связывающих домена одного димера контактируют с ДНК, связываясь с двумя последовательными витками большой бороздки. Такой тип взаимодействия в значительной степени повышает аффинность связывания репрессора с оператором. Ключевым элементом реакции взаимодействия является связывание короткой спирали шарнира с малой бороздкой ДНК последовательности оператора под углом ~45°. Такой изгиб ориентирует большую бороздку ДНК для взаимодействия с мотивом HTH.

Связывание индуктора вызывает немедленные конформационные изменения репрессора. Связывание двух молекул индуктора с тетрамерным репрессорным белком является достаточным условием для блокирования процесса репрессии. Это взаимодействие нарушает

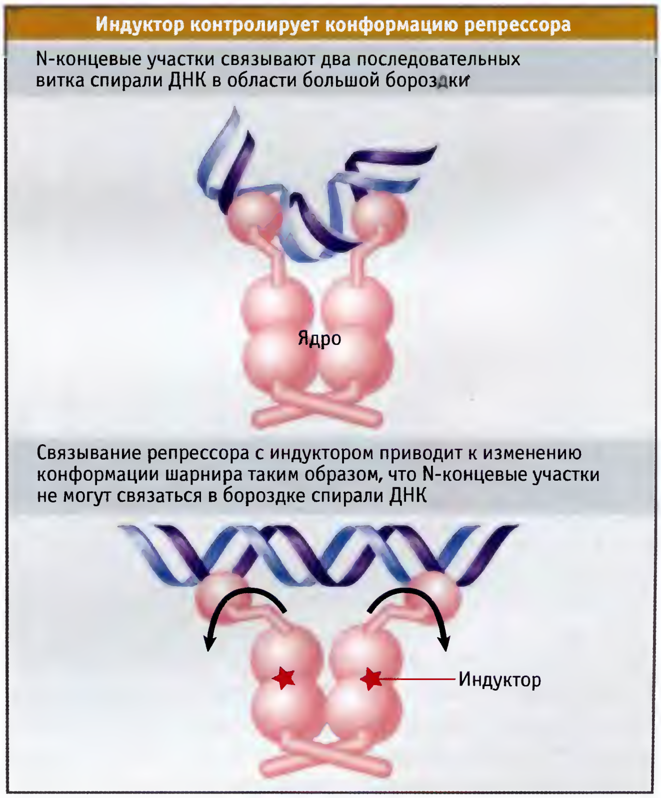


РИС. 12.15 Индуктор изменяет структуру ядра, нарушая спиральную структуру петли. При этом HTH-регионы теряют ориентацию, в которой способны правильно взаимодействовать с ДНК

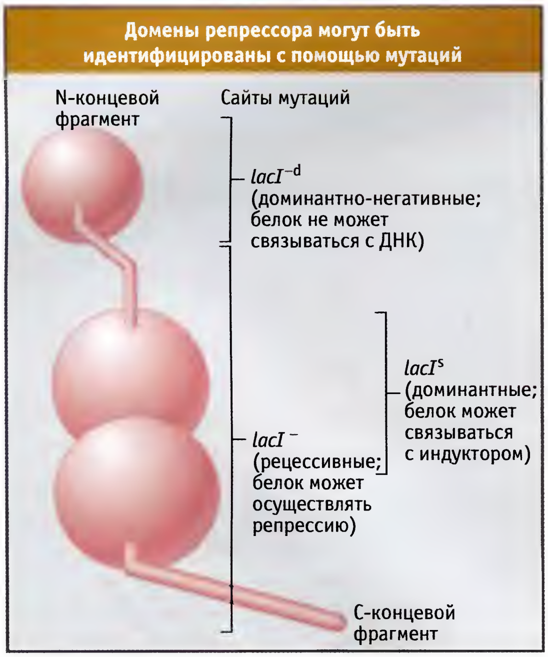
спиральную структуру петли и изменяет ориентацию головных фрагментов по отношению к ядру белка. В результате два N-концевых димера больше не могут связывать ДНК одновременно. Это явление сводит к минимуму все преимущества мультимерной структуры репрессорного белка и снижает аффинность его связывания с оператором.

12.13 Мутантные фенотипы коррелируют со структурой отдельных доменов репрессора

Ключевые понятия:

- Определенные типы мутаций свойственны отдельным доменам субъединиц репрессора.

Еще до выяснения точной структуры репрессорного белка с помощью мутаций было установлено, что он состоит из нескольких доменов. Сейчас на основе известной структуры мы можем более полно объяснить природу этих мутаций. Зависимость эффекта мутации от ее локализации схематично представлена на ·



12.16 Местоположение трех типов мутаций лактозного репрессора изображено на схеме доменной организации белка. Рецессивные *lacI*⁻ мутанты, неспособные к репрессии, могут обнаруживаться в любом участке молекулы. Доминантно-негативные мутанты *lacI*$^{-d}$, неспособные к репрессии, картируются в ДНК-связывающем участке. Доминантные *lacI*^{S} мутанты с нарушенным процессом индукции, в котором блокирована стадия связывания репрессора с индуктором или стадия аллостерического изменения конформации репрессора, обнаруживаются в ядерном домене 1

Рецессивные мутации типа *lacI*⁻ могут возникать в любом участке молекулы белка, за исключением концевых фрагментов. В целом любая мутация, инактивирующая белок, будет относиться к типу *lacI*⁻. Более детальное картирование мутаций на структуре кристалла, изображенного на рис. 12.13, позволяет определить точный эффект некоторых из них, к примеру влияние на процесс олигомеризации.

Особый класс доминантно негативных мутаций *lacI*^{d} обнаруживается в ДНК-связывающем сайте субъединиц репрессора (см. разд. 12.9 «Мультимерные белки обладают особыми генетическими свойствами»). Местоположение этих мутаций обусловливает нарушенную способность тетрамера, содержащего дефектные субъединицы, связываться с оператором. Снижение числа связывающих сайтов сокращает аффинность связывания репрессора с оператором. Роль N-концевого региона в специфическом связывании с ДНК подтверждается и обнаруживающимися в нем так называемыми мутациями «прочного связывания». Эти мутации повышают аффинность связывания репрессора с оператором, иногда до такой степени, что индуктор оказывается неспособным разрушить это взаимодействие. Однако такие мутации редки.

Неиндуцибельные мутации *lacI*^{S} расположены главным образом в регионе ядерного домена 1, находящегося между индуктор-связывающим сайтом и шарниром. Одна группа таких мутаций затрагивает аминокислоты, контактирующие с индуктором, препятствуя тем самым связыванию репрессора с индуктором. Другая группа свойственна сайтам, отвечающим за аллостерический контроль конформации шарнира, когда индуктор уже связан.

12.14 Репрессорный белок связывается с оператором

Ключевые понятия:

- Репрессорный белок связывается с двухцепочечной ДНК-последовательностью оператора.
- Оператор представляет собой палиндром длиной 26 пн.
- Каждый инвертированный повтор оператора связывается с ДНК-связывающим сайтом субъединицы репрессора.

Впервые репрессор был получен при выделении компонента, способного связываться с неметаболизируемым индуктором ИПТГ. (Поскольку количество репрессора в клетке слишком мало, для получения достаточного количества материала необходимо было использовать мутацию промотора, повышающую эффективность транскрипции гена *lacI*, и ввести такой мутантный локус в молекулу ДНК, присутствующую в клетке в большом числе копий. В результате удалось увеличить общую продукцию репрессора в 100–1000 раз.)

Репрессор связывается с двухцепочечной ДНК, содержащей последовательность *lac*-оператора дикого



РИС. 12.17 Последовательность нуклеотидов оператора *lac* симметрична. Последовательность пронумерована относительно стартовой точки транскрипции, обозначенной +1. Розовыми стрелками, направленными вправо и влево, показаны две группы повторов. Зелеными блоками отмечены идентичные основания

типа. Связывания репрессора не происходит, если ДНК выделена из мутанта O^c. При добавлении ИПТГ репрессор отделяется от оператора *in vitro*. Реакция между репрессором и оператором *in vitro* контролируется так же, как и *in vivo*, и, следовательно, может быть использована для изучения механизма репрессии.

Как репрессор узнает специфическую последовательность оператора на ДНК? Оператору присуще свойство, общее для многих сайтов узнавания бактериальными регуляторными белками: он представляет собой **палиндром**. Инвертированные повторы изображены на РИС. 12.17. На каждый повтор приходится половина связывающей области оператора.

Для определения точек контакта репрессора с оператором используют те же подходы, которые применялись для анализа взаимодействия промотора и полимеразы (см. разд. 11.14 «РНК-полимераза связывает молекулы ДНК с одной стороны»). Делеции генетического материала позволяют определить концевые точки области связывания; конститутивные точечные мутации дают возможность идентифицировать отдельные пары оснований, имеющие критическое значение. Были проведены эксперименты, в которых связанную и не связанную с репрессором ДНК сравнивали по чувствительности к метилированию или образованию поперечных сшивок в ДНК под действием УФ-облучения. В результате были выявлены основания, которые в связанном с белком состоянии либо лучше защищены, либо, наоборот, более чувствительны к обработкам, когда они связаны с белком.

Как показано на РИС. 12.18, участок ДНК, связанный с репрессором и защищенный тем самым от действия нуклеаз, расположен внутри области симметрии. На эту область приходится 26 пар оснований, находящихся между нуклеотидами −5 и +21. Зона, идентифицированная с помощью конститутивных мутаций, имеет меньший размер. В пределах центральной области, представленной 13 парами оснований и расположенной между нуклеотидами +5 и +17, содержится восемь сайтов, в которых замещение даже одной пары оснований приводит к конститутивности. В свете этих данных особое значение получают результаты изучения мутаций промотора, суммированные ранее на

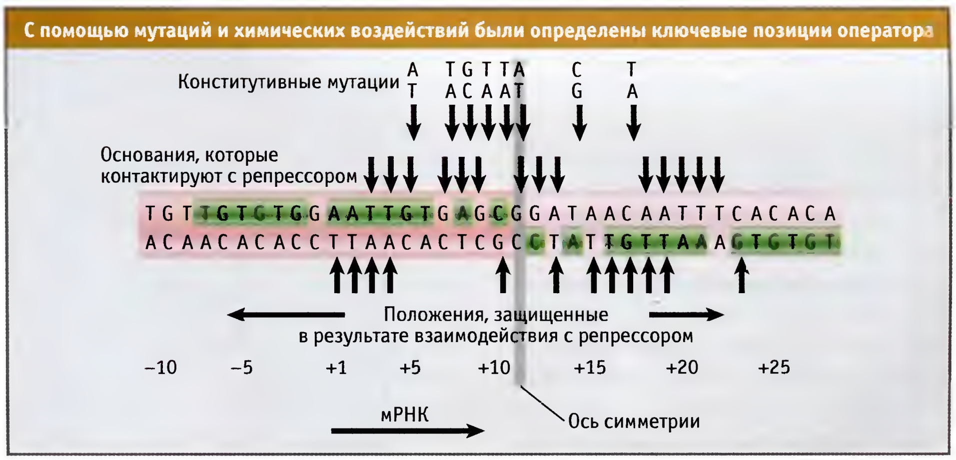


РИС. 12.18 Основания, взаимодействующие с репрессором, могут быть выявлены с помощью химических сшивок или с помощью модификаций, нарушающих процесс связывания. При использовании обоих подходов было установлено, что в связывании с репрессором участвует область, включающая в себя нуклеотиды от +1 до +23 на обеих цепях ДНК. Конститутивные мутации обнаруживаются в восьми позициях оператора (между нуклеотидами +5 и +17)

рис. 11.29. *За связывание специфических последовательностей ДНК с белком может быть ответственно небольшое число контактов в пределах большой по размеру области.*

Симметрия в последовательности ДНК отражает симметрию в белке. Каждая из идентичных друг другу субъединиц тетрамерного репрессора имеет ДНК-связывающий участок. Два таких связывающих сайта контактируют с оператором, а со стороны оператора в формировании контакта участвуют два инвертированных повтора. Такое взаимодействие проявляется в симметричности контактов, образующихся между репрессором и оператором (контакты между основаниями +1 и +6 являются симметричными по отношению к контактам между основаниями + 21 и +16), а также в распределении конститутивных мутаций внутри каждого инвертированного повтора. (Тем не менее оператор имеет нестрого симметричную структуру; его левая сторона связывается с репрессором прочнее, чем правая. Более сильный оператор можно создать путем дупликации (с инверсией) левой части и удалением центральной пары оснований.)

12.15 Связывание репрессора с индуктором приводит к его отделению от ДНК

Ключевые понятия:

- **Связывание репрессора с индуктором приводит к таким изменениям его конформации, в результате которых аффинность связывания с ДНК падает, и репрессор отделяется от оператора.**

Различные индукторы вызывают характерное уменьшение аффинности связывания репрессора с оператором *in vitro*. Эти изменения коррелируют с эффективностью индуктора *in vivo*. Следовательно, индукция основана на уменьшении сродства между репрессором и оператором. Проникая в клетку, индуктор связывается со свободными репрессорами, предотвращая их взаимодействие с операторами. Однако в том случае, когда репрессор уже прочно связан с оператором, как индуктор разрушает это взаимодействие?

На РИС. 12.19 приведены две модели, иллюстрирующие механизм работы индуктора:

- Модель равновесия (изображенная слева) предполагает, что репрессор, связанный с ДНК, находится в равновесии со свободным репрессором. Индуктор связывает свободную форму репрессора и тем самым нарушает равновесие путем предотвращения его повторного связывания с ДНК.
- Однако скорость диссоциации репрессора от оператора оказалась слишком низкой, чтобы соответствовать данной модели (период полураспада комплекса репрессор-оператор в отсутствие индуктора *in vitro* составляет >15 мин). Это означает, что *индуктор должен связываться непосредственно с репрессором, находящимся в комплексе с оператором.* Как следует из модели, представленной справа, присоединение индуктора должно приводить к таким изменениям репрессора, в результате которых он отделяется от оператора. И действительно, при добавлении ИПТГ *in vitro* происходит немедленная дестабилизация комплекса репрессор-оператор.

Связывание комплекса репрессор-ИПТГ с оператором может быть изучено при помощи больших концентраций белка в подходе с определением эффектов защиты/энхансинга при метилировании. Большое количество белка компенсирует низкую аффинность связывания комплекса репрессор-ИПТГ с оператором. Комплекс формирует с ДНК точно такие же контакты, как и свободный репрессор. Аналогичные результаты были получены в экспериментах с мутантными репрессорами, сродство которых к оператору, напротив, повышено; они образуют тот же тип контактов.

В целом типы репрессоров, у которых различия в аффинности к оператору достигают семи порядков, формируют с ДНК одни и те же контакты. *Следовательно, изменения в аффинности связывания репрессора с ДНК должны происходить путем влияния на общую конформацию ДНК-связывающего региона, а не путем создания или разрыва одной или нескольких связей.*

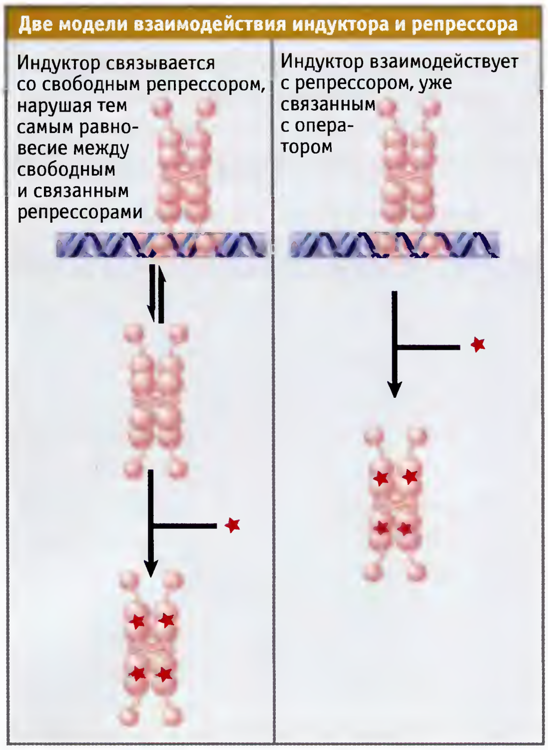


РИС. 12.19 Взаимодействует ли индуктор со свободным репрессором, нарушая равновесие (слева) или же с репрессором, уже связанным с оператором (справа)?

12.16 Репрессор способен связывать три оператора и взаимодействовать с РНК-полимеразой

Ключевые понятия:

- Каждый димер тетрамерного репрессора может связывать оператор, следовательно, тетрамер может связывать два оператора одновременно.
- Реализация полной репрессии требует, чтобы репрессор провзаимодействовал с дополнительным оператором, расположенным выше или ниже основного оператора, перекрывающегося с промотором гена *lacZ*.
- Связывание репрессора с оператором способствует связыванию РНК-полимеразы с промотором, однако препятствует транскрипции.

Эффект аллостерического изменения конформации при связывании репрессора с индуктором проявляется уже в димере. Почему, в таком случае, для реализации полной репрессии необходимо, чтобы репрессор имел тетрамерную структуру?

Каждый димер может связываться с последовательностью оператора. Это означает, что интактный репрессор способен связать два оператора одновременно. В действительности, в стартовом участке оперона *lac* расположены две последовательных операторных сайтах. Первый оператор, *O1*, расположен непосредственно перед началом гена *lacZ* и характеризуется высокой аффинностью к репрессору. Еще два оператора (иногда называемые псевдо-операторами) характеризуются пониженным сродством к репрессору; *O2* расположен на 410 оснований ниже стартовой точки гена *lacZ*, а *O3* — на 88 оснований выше стартовой точки гена *lacI*.

На РИС. 12.20 показано, что происходит в том случае, когда ДНК-связывающий белок связывается с двумя отдельными сайтами ДНК одновременно. Последовательность ДНК между этими сайтами формирует петлю. Длина петли зависит от расстояния между двумя сайтами связывания. Когда Lac-репрессор связывается одновременно с *O1* и с одним из двух других операторов, ДНК между этими участками формирует короткую петлю, сжимающую структуру ДНК. На РИС. 12.21 представлена модель взаимодействия тетрамерного репрессора с двумя операторами.

Связывание с дополнительными операторами влияет на уровень репрессии. Удаление какой-либо из последовательностей *O2* или *O3* приводит к снижению эффективности репрессии в 2–4 раза, удаление обеих последовательностей — в 100 раз. *Из этого следует, что способность репрессора связываться наряду с O1 с одним из двух других операторов важна для реализации репрессии.* Эксперименты *in vitro* со сверхспирализованными плазмидами, содержащими множественные операторы, выявили существенную стабилизацию lacI–ДНК комплекса. Тем не менее при связывании lacI с ИПТГ репрессор быстро диссоциировал от ДНК.

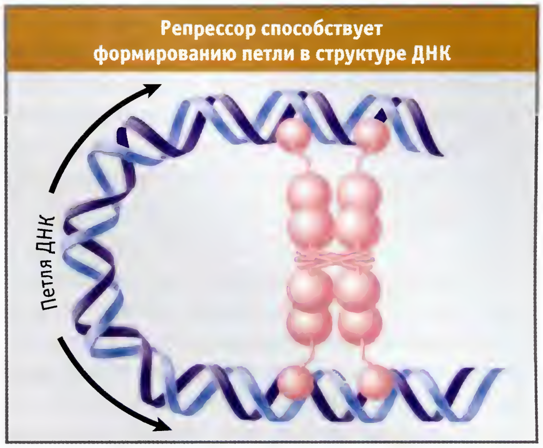


РИС. 12.20 Когда оба димера тетрамерного репрессора связываются с ДНК, ДНК формирует петлю



РИС. 12.21 Когда тетрамерный репрессор связывается с двумя операторами, последовательность ДНК между участками связывания сворачивается в жесткую петлю. (Структура синего цвета в центре свернувшейся ДНК — CAP, еще один регуляторный белок, связывающийся в этом регионе.) Данные воспроизведены с разрешения Lewis M., et al. 1996. *Science*. 271: cover. 1996 AAAS. Любезно предоставлено Ponzy Lu, University of Pennsylvania

Кроме того, мы знаем о прямых последствиях связывания репрессора с оператором (*O1*). Первоначально считалось, что это взаимодействие препятствует контакту РНК-полимеразы с промотором. Сейчас мы знаем, что два белка могут связываться с ДНК одновременно, при этом *связывание ДНК с репрессором увеличивает прочность связывания ДНК с РНК-полимеразой!* Тем не менее связанный фермент не может инициировать транскрипцию.

Константа равновесия для реакции связывания РНК-полимеразы с *lac*-промотором составляет $1{,}9 \times 10^7$ М$^{-1}$. В присутствии репрессора эта константа увеличивается на два порядка и составляет $2{,}5 \times 10^9$ М$^{-1}$. Если ориенти-

роваться на значения константы равновесия K_B, приведенные на рис. 11.19, можно сказать, что репрессор превращает процесс формирования закрытого комплекса между РНК-полимеразой и *lac*-промотором из слабого взаимодействия в прочное.

Какое последствие это имеет для индукции оперона? Высокое значение K_B свидетельствует о том, что занятый репрессором промотор в 100 раз чаще останется связанным с РНК-полимеразой. Кроме того, одновременное взаимодействие ДНК с репрессором и РНК-полимеразой делает возможным немедленное начало транскрипции сразу после индукции. При этом нет необходимости ждать, пока РНК полимераза сядет на промотор.

Иными словами, репрессор обеспечивает «хранение» РНК-полимеразы на промоторе. Комплекс РНК-полимераза–репрессор–ДНК блокирован в закрытом состоянии. При добавлении индуктора репрессор высвобождается, и закрытый комплекс превращается в открытый, что и инициирует транскрипцию. Общий эффект взаимодействия с репрессором проявляется в ускорении процесса индукции.

Применима ли эта модель к другим системам? Взаимодействие между РНК-полимеразой, репрессором и промоторными/операторными участками на ДНК характеризуется своими особенностями для каждой системы. Обусловлено это тем, что последовательность оператора не всегда перекрывается с промотором (см. рис. 12.26). К примеру, у фага лямбда оператор расположен выше промоторного участка, и связывание с репрессором блокирует связывание с РНК-полимеразой (см. гл. 14 «Стратегии фагов»). То есть взаимодействие репрессора, связанного с ДНК, и РНК-полимеразы свойственно не всем системам.

12.17 Репрессор всегда связан с ДНК

Ключевые понятия:

- **Белки, обладающие высокой аффинностью к определенным последовательностям ДНК, в отношении всех остальных последовательностей проявляют низкую аффинность.**
- **Каждая пара оснований бактериального генома может стать началом низкоаффинного сайта связывания для репрессора.**
- **Наличие большого количества низкоаффинных сайтов связывания гарантирует, что все молекулы репрессорного белка будут ассоциированы с ДНК.**
- **Репрессор находит оператор, перемещаясь к нему по ДНК от случайного сайта связывания, а не диффундируя из раствора.**

Вероятно, все белки, проявляющие большую специфичность в отношении определенной последовательности ДНК, проявляют низкое сродство в отношении каких-либо случайных последовательностей. Большое число низкоаффинных сайтов связывания будет конкурировать с небольшим числом специфичных участков за связывание с репрессором-тетрамером. В геноме *E.coli* существует только один сайт, обладающий высокой аффинностью к репрессору, — оператор. Все остальные последовательности — это сайты связывания, характеризующиеся низким сродством к репрессору. Каждая пара оснований бактериального генома может стать началом низкоаффинного сайта связывания для репрессора. (Для формирования низкоаффинного сайта связывания достаточно смещения репрессора вдоль последовательности ДНК всего на одно основание за пределы оператора!) В результате количество таких случайных сайтов связывания составляет $4{,}2 \times 10^6$.

Большой избыток случайных участков связывания означает, что даже в отсутствие участка специфического связывания, все или почти все молекулы репрессора будут связаныс ДНК. В растворе (в свободном состоянии) репрессор обнаруживаться не будет. На РИС. 12.22 приведено уравнение, описывающее равновесие между свободным репрессором и репрессором, связанным с ДНК. Преобразовав это уравнение, можно рассчитать долю свободного репрессора от общего количества репрессора в клетке.

Используя в расчетах параметры системы *lac*, определяем, что:

- Константа равновесия для реакции неспецифического связывания равна $K_A = 2 \times 10^6\ М^{-1}$.
- Концентрация неспецифических последовательностей ДНК составляет 4×10^6 на 10^{-15} л бактериального объема, что соответствует концентрации [ДНК] = 7×10^{-3} М (очень высокая концентрация).

Выражая одну величину через другую, получаем, что соотношение свободного репрессора к связанному составляет 10^{-4}.

За исключением 0,01%, все молекулы репрессора связаны со случайными последовательностями ДНК. А поскольку на одну клетку приходится ~ 10 молекул репрессора, можно утверждать, что свободного репрессора в клетке вовсе нет. Этот факт вносит значимый вклад в понимание механизмов взаимодействия репрессора с оператором: один-единственный высокоспецифичный сайт конкурирует с большим количеством случайных участков за связывание с репрессором.

При такой конкуренции абсолютные величины констант ассоциации репрессора с оператором и репрессора со случайной последовательностью значения не имеют.

Репрессор+ДНК ⇌ репрессор-ДНК

Равновесие реакции связывания репрессора со случайной последовательностью ДНК описывается уравнением:

$$K_A = \frac{[\text{Репрессор-ДНК}]}{[\text{Свободный репрессор}]\,[\text{ДНК}]}$$

Долю свободного репрессора можно рассчитать, преобразовав уравнение:

$$\frac{[\text{Свободный репрессор}]}{[\text{Репрессор-ДНК}]} = \frac{1}{K_A \times [\text{ДНК}]}$$

РИС. 12.22 **Реакция связывания репрессора со случайной последовательностью описывается уравнением равновесия**

В этом случае существенным параметром становится отношение K_{sp} (константа реакции связывания со специфическим сайтом) к K_{nsp} (константа реакции связывания со случайной последовательностью), то есть специфичность.

12.18 Оператор конкурирует со случайными последовательностями за связывание с репрессором

Ключевые понятия:

- **В отсутствие индуктора аффинность репрессора к оператору превышает его сродство к случайной последовательности в 10^7 раз.**
- **Количество молекул репрессора на клетку (около 10) — достаточно для того, чтобы обеспечить связывание оператора в течение 96% времени.**
- **При добавлении индуктора разница между аффинностью репрессора к оператору и к неспецифическим сайтам сокращается с 10^7 до 10^4. При этом связывается только 3% операторов.**
- **Присоединение индуктора приводит к смещению репрессора с последовательности оператора на случайную последовательность путем прямой замены одного сайта на другой.**
- **Эти параметры могут быть изменены путем повышения или понижения эффективной концентрации ДНК *in vivo*.**

Мы можем определить параметры, влияющие на способность регуляторных белков связываться с последовательностями-мишенями, сравнивая уравнения равновесия для реакций специфического и неспецифического связывания. Перечислим важные параметры:

- Размер генома влияет на способность белка связываться со специфической последовательностью.
- Реакция связывания с ДНК зависит, с одной стороны, от специфичности белка, а с другой — от общей массы ДНК.
- Количество белка, требуемое для реализации специфичного связывания, увеличивается с ростом генома и уменьшается с повышением специфичности белка.
- Количество белка должно находиться в избытке по отношению к общему числу специфических последовательностей-мишеней. В связи с этим регуляторных белков, способных связываться с большим количеством мишеней, обнаружено гораздо больше, чем тех, которые реагируют с ограниченным числом мишеней.

На РИС. 12.23 приведено сравнение констант равновесия для реакций связывания репрессор/оператор и репрессор/случайная последовательность ДНК. Основываясь на этих значениях, мы можем предположить, как репрессор распределен между оператором и неспецифическими сайтами связывания на ДНК, а также каким образом присоединение индуктора приводит к диссоциации репрессора от оператора.

К оператору репрессор присоединяется ~ в 10^7 раз лучше, чем к любой другой случайной последовательности ДНК той же длины. Оператор представляет собой единственный высокоаффинный сайт, конкурирующий с большим количеством случайных последовательностей ДНК за связывание с репрессором. Каким же образом в этом случае репрессор обеспечивает эффективный контроль экспрессии оперона?

Основываясь на специфичности белка, мы можем оценить распределение репрессора между оператором и случайными сайтами связывания. При общем количестве молекул *lac*-репрессора на клетку (равном 10) и специфичности его связывания с оператором 10^7, можно подсчитать, что оператор будет связан с репрессором 96% времени. Роль специфичности объясняют две особенности взаимодействия между репрессором и оператором:

- Присоединение индуктора к репрессору приводит к уменьшению аффинности репрессора к оператору ~ на 10^3. Сродство же в отношении случайных последовательностей остается неизменным. В результате специфичность связывания репрессора с оператором снижается до 10^4. Этой величины уже недостаточно для того, чтобы обеспечить предпочтительное связывание репрессора с оператором, а не с низкоаффинными сайтами, присутствующими в избытке ($4{,}2 \times 10^6$). В таких условиях только 3% операторов будут связываться репрессором.
- Мутации, уменьшающие сродство репрессора к оператору в 20–30 раз, приводят к конститутивной экспрессии генов оперона. В пределах генома мутантные сайты могут утратить преимущество в отношении случайных сайтов. При уменьшении специфичности репрессора в 10 раз эффективность связывания репрессором оператора снижается примерно на 50%.

В результате такого взаимодействия в неиндуцированных клетках одна молекула репрессора-тетрамера присоединяется к оператору. Все или почти все оставшиеся тетрамеры будут присоединяться случайно к другим

Репрессор специфически связывается с оператором

ДНК	Репрессор	Репрессор + индуктор
Оператор	2×10^{13}	2×10^{10}
Случайная последовательность	2×10^6	2×10^6
Специфичность	10^7	10^4
Доля связанных операторов	96%	3%
Состояние оперона	репрессирован	индуцирован

РИС. 12.23 Lac-репрессор прочно и специфично связывается с оператором. Однако присоединение индуктора приводит к отделение репрессора. Все константы равновесия приведены в M^{-1}

последовательностям ДНК, как показано на РИС. 12.24. Свободные тетрамеры репрессора, по-видимому, либо не содержатся в клетке, либо их очень мало.

Присоединение индуктора нарушает способность репрессора специфично связываться с оператором. Те же репрессоры, что уже связаны с оператором, диссоциируют от него и взаимодействуют со случайными (низко аффинными) последовательностями. В результате в индуцированной клетке тетрамеры репрессора будут «храниться» в случайных сайтах ДНК. В неиндуцированной клетке один тетрамер будет присоединяться к оператору, в то время как остальные молекулы репрессора будут присоединяться к неспецифичным сайтам. *Следовательно, эффект индукции сводится не к появлению свободных молекул репрессора, а к изменению его распределения на ДНК.*

При удалении индуктора репрессор очень быстро восстанавливает свою способность специфично взаимодействовать с оператором. Это происходит в результате его смещения с неспецифических сайтов «хранения». Какими механизмами обеспечивается быстрый поиск оператора? Способность репрессора очень быстро связываться с оператором не согласуется со временем, требуемым для множественных диссоциаций и реассоциаций комплекса репрессор-ДНК. Это исключает наличие механизма беспорядочного поиска оператора и дает основание полагать, что репрессор способен передвигаться непосредственно от случайного сайта к оператору. Этот процесс происходит с помощью того же механизма, который, как мы уже обсуждали, РНК-полимераза использует для поиска промотора (см. рис. 11.22 и 11.23). Движение может осуществляться за счет быстрого скольжения вдоль ДНК либо за счет непосредственного перемещения из сайта в сайт (как показано на рис. 11.24). Реакцию перемещения может облегчить наличие большего числа (четырех) сайтов связывания (в расчете на тетрамер), чем на самом деле используется для контактов с ДНК в каждый момент времени (два из четырех).

В условиях конкуренции между единственным высокоаффинным оператором и большим числом низкоаффинных случайных последовательностей репрессор встает перед дилеммой. Для реализации репрессии он должен быть высокоспецифичен к оператору. Для реализации же индукции эта специфичность должна быть снижена. Предположим, к примеру, что на клетку приходится 1000 молекул репрессора: тогда в условиях репрессии лишь 0,04% операторов будут свободны, и только 40% операторов будут свободны в условиях индукции.

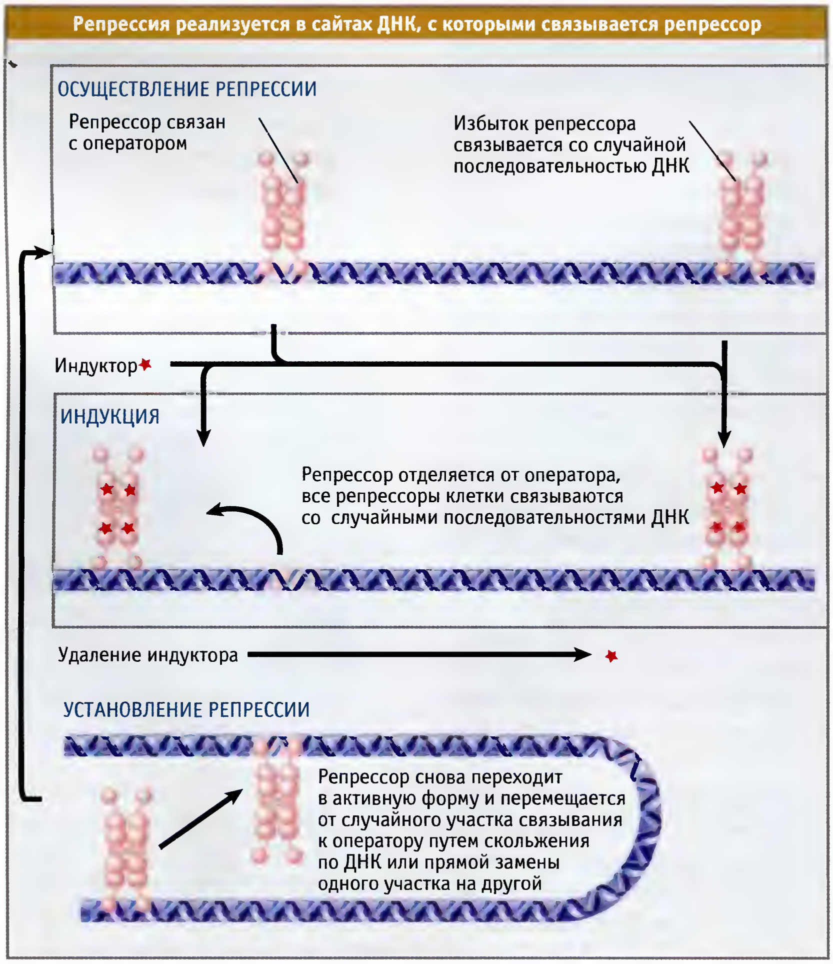


РИС. 12.24 Все репрессоры клетки связаны с ДНК

Мы видим обратную корреляцию между способностью репрессора достигать полной репрессии оператора и способностью эффективно снижать специфичность связывания во время индукции. Основываясь на этом, можно предположить, что число молекул репрессора, синтезируемых *in vivo*, отражает баланс между этими двумя состояниями.

Разница в уровне экспрессии лактозного оперона в его индуцированном и репрессированном состоянии *in vivo* составляет 10^3 раз. Иными словами, даже в отсутствие индуктора существует определенный фоновый уровень экспрессии, составляющий 0,1% от индуцированного уровня. Он может быть ниже — в присутствии большего количества белка-репрессора — или выше — если количество белка уменьшено. Так, состояние полной репрессии не может быть достигнуто, если в клетке присутствует менее 10 молекул репрессора; и могут возникнуть трудности при индукции оперона, если молекул репрессора слишком много. Были проведены эксперименты, в ходе которых систему *lac*-репрессор-оператор вводили мышам. Когда оператор *lac* связан с репортерным геном тирозиназы, продукция фермента индуцируется добавлением ИПТГ. Это означает, что для поиска в геноме мыши своей мишени репрессору потребуется времени в 10^3 раз больше, чем в геноме *E. coli*. При этом индукция осуществляется примерно теми же концентрациями ИПТГ, что и индукция в бактериальной клетке. Эффективности, с которой происходит индукция, а также требуемых концентраций Lac-репрессора мы, тем не менее, не знаем.

Для того чтобы данные, полученные при изучении аффинности ДНК-белковых взаимодействий *in vitro*, экстраполировать на процессы, протекающие *in vivo*, нам необходимо знать эффективные концентрации ДНК *in vivo*. Значение «эффективной концентрации» отличается от отношения масса/объем из-за нескольких факторов. Эффективная концентрация увеличивается при молекулярной агрегации, когда поливалентные катионы нейтрализуют ~90% зарядов на ДНК, и нуклеиновые кислоты конденсируются с образованием макромолекулярных структур. Главным фактором, снижающим эффективную концентрацию, является недоступность ДНК вследствие ее конденсации или связывания с белками.

Одним из способов определить эффективную концентрацию является сравнение зависящей от концентрации ДНК скорости реакции *in vitro* и *in vivo*. Это было сделано с использованием межмолекулярной рекомбинации двух молекул ДНК. В качестве контроля проводили реакцию внутримолекулярной рекомбинации, то есть реакцию рекомбинации между сайтами, расположенными в пределах одной молекулы ДНК. Мы предполагаем, что *межмолекулярная* реакция *in vivo* и *in vitro* протекает при одних и тех же концентрациях ДНК, а потому любая разница в скорости реакции внутри-/межмолекулярной рекомбинации будет обусловлена изменением эффективной концентрации ДНК *in vivo*. Результаты таких экспериментов позволяют предположить, что эффективная концентрация ДНК *in vivo* снижается более чем в 10 раз.

Реакции рекомбинации ДНК, а также ДНК-белковых взаимодействий зависят от концентрации ДНК, а следовательно, чувствительны к процессам, приводящим к ее снижению. Эта проблема встает перед всеми ДНК-связывающими белками в процессе быстрого поиска ими мишени. Кроме того, существование такого явления исключает возможность поиска мишеней для связывания путем случайной диффузии (см. рис. 11.22).

12.19 Репрессия может происходить сразу в нескольких локусах

Ключевые понятия:

- Репрессор может действовать на все участки ДНК, последовательность которых совпадает с последовательностью оператора.

Репрессор Lac действует только на оператор генов кластера *lacZYA*. Но некоторые репрессоры могут контролировать диффузно расположенные структурные гены, взаимодействуя сразу с несколькими операторами. Примером может служить репрессор *trp*, контролирующий три физически не связанных группы генов:

- Оператор кластера структурных генов *trpEDBCA* контролирует координированный синтез ферментов, участвующих в синтезе триптофана из хоризмовой кислоты.
- Оператор другого локуса контролирует экспрессию гена *aroH*. Этот ген кодирует один из трех ферментов, катализирующих начальную стадию в общем пути биосинтеза ароматических аминокислот.
- Регуляторный ген *trpR* репрессируется своим же собственным продуктом — *trp*-репрессором. Следовательно, репрессорный белок действует, снижая и свой собственный синтез. Такой цикл может служить примером **аутогенного** контроля (самоконтроля). Подобное явление весьма распространено в регуляторных генах, этим способом может осуществляться как негативный, так и позитивный контроль (см. разд. 12.23 «Синтез r-белка контролируется с помощью аутогенной регуляции» и разд. 14.15 «Репрессор поддерживает цикл аутогенной регуляции»).

Родственная последовательность оператора протяженностью 21 пн имеется в каждом из трех локусов, в которых действует *trp*-репрессор. Степень консервативности этих последовательностей приведена на РИС. 12.25. Каждому оператору свойственна существенная (но не идентичная) двойная симметрия. Консервативные особенности структуры, свойственные для всех трех операторов, указывают на точки, существенные для контакта с *trp*-репрессором. Это объясняет, как один репрессорный белок действует в нескольких локусах: *каждый локус*



РИС. 12.25 Репрессор *trp* распознает операторы сразу в трех локусах. Консервативные нуклеотиды показаны красным цветом. Положение стартовой точки варьирует, что показано с помощью стрелок

имеет копию специфической нуклеотидной последовательности, узнаваемой репрессором (точно так же, как каждый промотор содержит консенсусную последовательность, характерную для других промоторов).

На РИС. 12.26 приведены различные варианты взаимного расположения операторов и промоторов. Отличительной чертой таких «разбросанных» операторов является их различная локализация относительно промотора в каждом локусе. В гене *trpR* оператор расположен между позициями –12 и +9, тогда как в *trp*-опероне он лежит между нуклеотидами –23 и –3. В локусе *aroH* оператор находится дальше от начала транскрипции, располагаясь между нуклеотидами –49 и –29. В других случаях оператор расположен после промотора (как в *lac*-системе) или прямо перед ним (как в *gal*-системе, для которой природа эффекта репрессии до конца не выяснена). Способность репрессоров действовать на операторы, различно расположенные в пределах промоторов, указывает на то, что в каждом конкретном случае репрессия происходит по своим механизмам. Однако общей чертой процесса является блокирование инициации транскрипции РНК-полимеразой на промоторе.

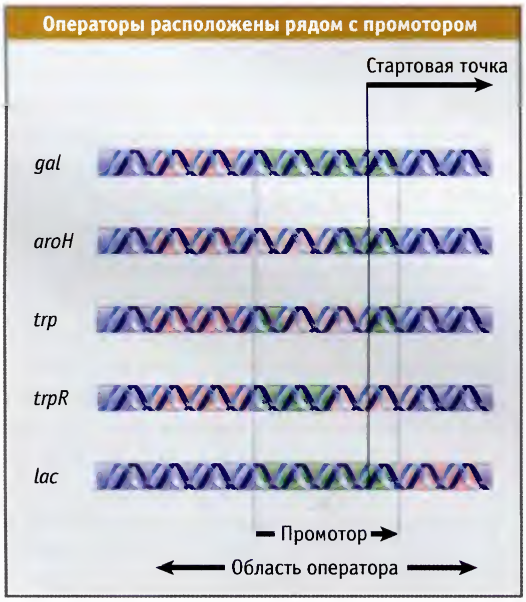


РИС. 12.26 Операторы могут располагаться в различных положениях относительно промотора

12.20 Цикло-АМФ — эффектор, активирующий БРЦ и опосредующий его действие на многие опероны

Ключевые понятия:

- БРЦ — активаторный белок, необходимый для инициации транскрипции на зависимых от него промоторах.
- Для активации димера БРЦ необходима одна молекула цАМФ.

До сих пор мы говорили о промоторах как о последовательностях ДНК, которые должны быть узнаны РНК-полимеразой для того, чтобы впоследствии фермент смог начать транскрипцию. Однако существуют промоторы, на которых РНК-полимераза не может начать транскрипцию без помощи дополнительных регуляторных белков. Такие белки служат позитивными регуляторами, присутствие которых необходимо для переключения транскрипционных единиц. Как правило, активатор компенсирует слабость промотора, к примеру недостаток консенсусных последовательностей –35 и –10.

Одним из наиболее часто встречающихся примеров таких активаторов служит белок, названный **БРЦ апоиндуктор** (белок-рецептор цАМФ, CRP), который контролирует большую группу оперонов в клетках *E. coli*. Этот белок является фактором позитивного контроля, его присутствие необходимо для инициации транскрипции на зависимых от него промоторах. БРЦ активен только

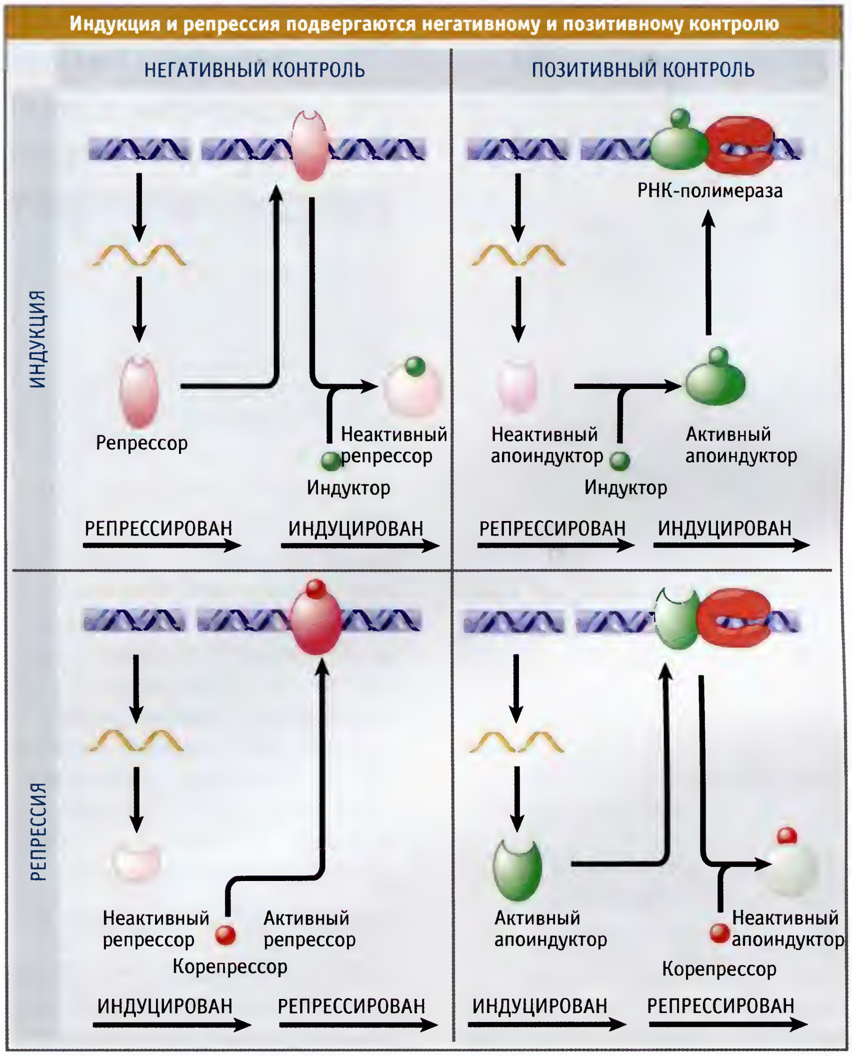


РИС. 12.27 Контролирующие системы отличаются гибкостью и предназначены для осуществления негативного или позитивного контроля индукции или репрессии

в присутствии цАМФ, который ведет себя как классический низкомолекулярный индуктор (см. РИС. 12.27; верхняя правая часть).

цАМФ синтезируется ферментом — **аденилатциклазой**. Субстратом для реакции служит АТФ, в который с помощью фосфодиэфирной группы вводится связь между 3′ и 5′-атомами углерода, в результате образуется структура, изображенная на РИС. 12.28. Мутанты по гену, кодирующему аденилатциклазу (*cya*⁻), не отвечают на изменения уровня глюкозы в клетке.

Уровень цАМФ обратно пропорционален уровню глюкозы. В основе этого эффекта лежат те же компоненты Pts-системы, которые ответственны за контроль потребления лактозы. Фосфорилированная форма белка IIA^{Glc} стимулирует аденилатциклазу. При поступлении в клетку глюкозы дефосфорилирование IIA^{Glc} приводит к снижению активности фермента.

Как показано на РИС. 12.29, низкий уровень цАМФ приводит к неспособности белка дикого типа связываться с контролируемым участком, что, в свою очередь, проявляется в неспособности РНК-полимеразы инициировать транскрипцию. Следовательно, снижая уровень цАМФ, глюкоза лишает соответствующие опероны регуляторного фактора, необходимого для их экспрессии.

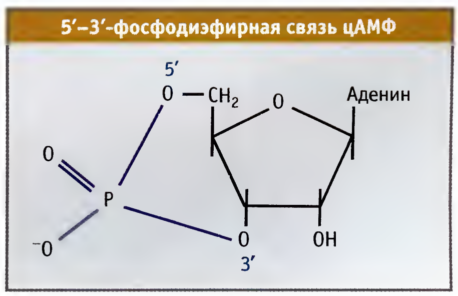
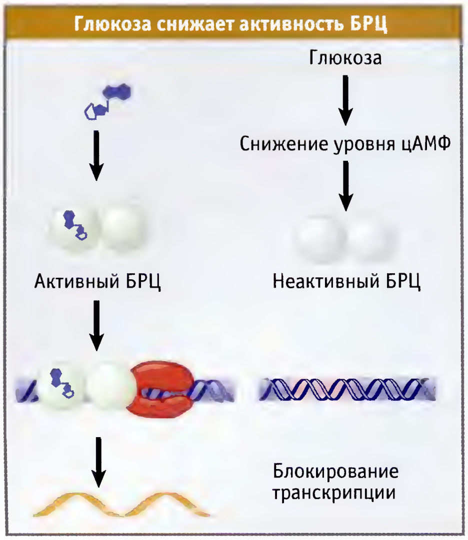


РИС. 12.28 Фосфатная группа циклического АМФ соединена с 3′- и 5′-атомами углерода рибозного кольца



Снижая уровень цАМФ, глюкоза ингибирует транскрипцию БРЦ-зависимых оперонов

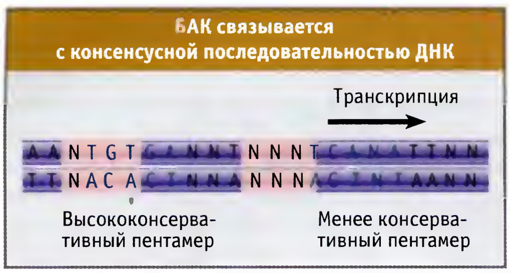


РИС. 12.30 БРЦ-связывающая область содержит в своем составе высококонсервативный пентамер TGTGA и (иногда) инвертированный повтор этой последовательности (TCANA)

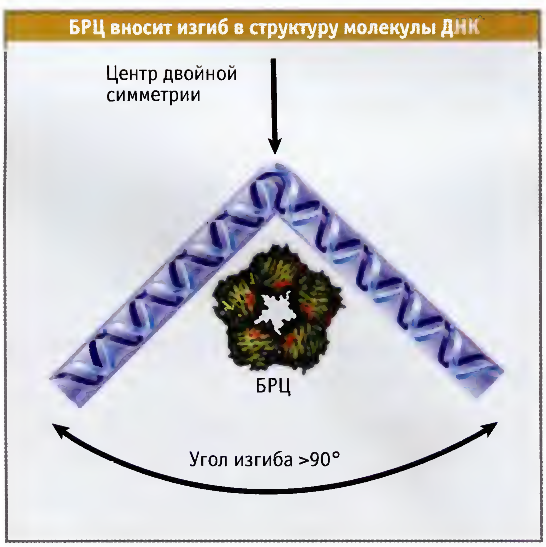


РИС. 12.31 Связывание с БРЦ приводит к изгибу молекулы ДНК вокруг центра симметрии больше чем на 90°

12.21 БРЦ функционирует на основе различных механизмов, действуя при этом на разные опероны

Ключевые понятия:

- Связываясь с БРЦ, ДНК изгибается на 90°.
- Сайты связывания с БРЦ в разных оперонах-мишенях характеризуются различной локализацией относительно промотора.
- БРЦ взаимодействует с РНК-полимеразой, однако детали этого процесса зависят от положения сайта связывания БРЦ относительно промотора.

Фактор БРЦ связывается с ДНК, и комплекс АМФ-БРЦ-ДНК можно обнаружить на каждом БРЦ-зависимом промоторе. Фактор представляет собой димер, состоящий из двух идентичных субъединиц с молекулярной массой 22,5 кДа. Для активации БРЦ необходима одна молекула цАМФ. Мономер БРЦ содержит в своем составе ДНК-связывающий участок и участок, отвественный за активацию транскрипции.

Димер БРЦ связывает участок ДНК, протяженностью ~22 пн. Возможные вариации в консенсусной последовательности сайта связывания приведены на РИС. 12.30. Мутации, препятствующие осуществлению БРЦ своей функции, как правило, затрагивают высококонсервативную пентамерную последовательность $\frac{\text{TGTGA}}{\text{ACACT}}$, которая, судя по всему, обеспечивает узнавание белком участка-мишени. Прочнее всего БРЦ связывает сайты, содержащие два (инвертированных) пентамерных повтора, поскольку это дает возможность связаться с ДНК обеим субъединицам димера. Во многих сайтах связывания второй пентамер отсутствует, однако связаться с ДНК все равно должны обе субъединицы, вторая при этом связывается со случайной последовательностью (если она вообще способна связываться с ДНК). Иерархия последовательностей согласно их аффинности связывания с БРЦ помогает объяснить, почему разным генам для активации *in vivo* необходимо разное количество цАМФ.

Связываясь с ДНК, БРЦ вносит в ее структуру существенный изгиб. В промоторе *lac* точка перегиба расположена в центре двойной симметрии. Как демонстрирует модель, приведенная на РИС. 12.31, молекула сгибается под достаточно большим углом (>90°). Организация двойной спирали ДНК при этом претерпевает существенные изменения. Резкий изгиб возникает в области консенсусной последовательности TGTGA. При связывании БРЦ с двумя инвертированными повторами консенсусной последовательности общий угол изгиба составляет 90°. Возможно, такой изгиб молекулы напрямую влияет на процесс транскрипции, однако он может быть необходим только для обеспечения контакта БРЦ и РНК-полимеразы в промоторе.

Любопытная особенность механизма действия БРЦ заключается в том, что его сайты-мишени в различных

оперонах имеют разную локализацию относительно стартовой точки транскрипции. Пентамер TGTGA может располагаться в любой ориентации. На РИС. 12.32 приведены три примера, описывающие местоположение возможных участков связывания БРЦ:

- БРЦ-связывающий сайт примыкает к промотору, как в *lac*-опероне, в котором центр участка, экранируемого БРЦ, находится в позиции –61. Возможно, в *lac*-опероне связываются два димера БРЦ. Связавшийся БРЦ расположен главным образом с одной стороны молекулы ДНК, с той же, с которой связывается и РНК-полимераза. Такой характер связывания дает белкам возможность беспрепятственно взаимодействовать друг с другом.
- Иногда БРЦ-связывающий сайт расположен внутри промотора, как в локусе *gal*, в котором центр участка связывания БРЦ находится в позиции –41. Вероятно, в этом регионе связывается только один димер БРЦ. Поскольку БРЦ-связывающий регион частично лежит в области промотора, обычно экранируемой РНК-полимеразой, можно предположить, что белки находятся в очень тесном контакте.
- В других оперонах БРЦ-связывающий сайт расположен перед промотором. Центр участка связывания с единственным димером БРЦ в регионе *ara* расположен дальше всего от стартовой точки транскрипции — в позиции –92.

Степень зависимости от регуляторного белка БРЦ определяется собственной эффективностью промотора. Промоторы, независимые от БРЦ, обладают высококонсервативной –35 последовательностью и в меньшей степени, но тем не менее консервативной –10 последовательностью. Действительно, промоторы с эффективными –35 и –10 регионами могут взаимодействовать с РНК-полимеразой независимо от регуляторных факторов. Контроль таких промоторов с помощью БРЦ трудно осуществим.

Различают два принципиально разных механизма, посредством которых БРЦ может активировать транскрипцию: он может напрямую взаимодействовать с РНК-полимеразой или он может, действуя на ДНК, менять ее структуру таким образом, чтобы облегчать связывание с ферментом. В действительности же БРЦ взаимодействует как с РНК-полимеразой, так и с ДНК.

Сайты связывания с БРЦ во многих промоторах напоминают либо БРЦ-связывающий сайт *lac*-оперона (центр последовательности расположен в положении –61), либо БРЦ-связывающий сайт *gal*-оперона (центр последовательности расположен в положении –41). Основное различие между ними заключается в том, что сайты связывания первого типа (названные класс I) расположены выше промотора, тогда как БРЦ-связывающие сайты второго типа (класс II) частично перекрываются с участком связывания РНК-полимеразы. (Характер взаимодействий в *ara*-опероне может быть различным.)

И в том, и в другом случае центр БРЦ-связывающей последовательности отстоит от стартовой точки транскрипции на целое число витков двойной спирали ДНК. Из этого следует, что БРЦ связывается на той же стороне молекулы ДНК, что и РНК-полимераза. Однако природа взаимодействий между БРЦ и РНК-полимеразой для двух типов промоторов различается.

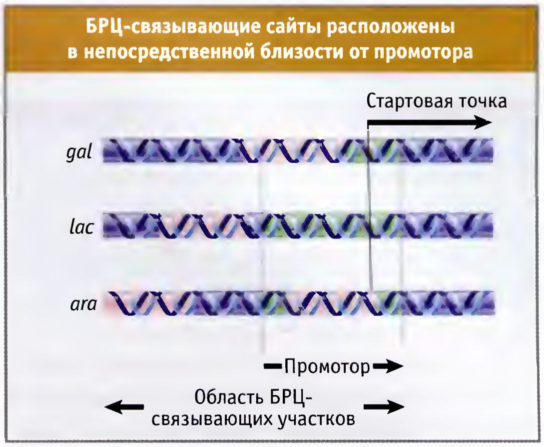


РИС. 12.32 Сайты-мишени БРЦ в различных оперонах имеют разную локализацию относительно стартовой точки транскрипции

Делеция в С-концевой части α-субъединицы РНК-полимеразы приводит к потере восприимчивости транскрипции к активации белком БРЦ. БРЦ имеет «активирующий домен», необходимый для активации промоторов обоих типов. Этот активирующий домен представляет собой петлю, состоящую из ~ 10 аминокислот. Он напрямую взаимодействует с α-субъединицей РНК-полимеразы, активируя фермент. Для активации транскрипции на промоторах I класса этого взаимодействия достаточно. Промоторы II класса требуют еще одного события, в котором участвуют другой регион БРЦ и N-концевой фрагмент α-субъединицы РНК-полимеразы.

Эксперименты с димерами БРЦ, в которых только одна субъединица обладала активирующим доменом, показали, что непосредственно в активации транскрипции участвует одна субъединица белка, расположенная ближе к стартовой точке и поэтому контактирующая с РНК-полимеразой. Объяснение этого явления может быть следующим: димерная структура БРЦ дает гарантию того, что одна субъединица будет всегда доступна для взаимодействия с РНК-полимеразой, независимо от того, какая субъединица связалась с ДНК и в какой ориентации.

Эффект, оказываемый регуляторным белком БРЦ на РНК-полимеразу, зависит от взаимного расположения двух белков. В промоторах I класса, где участок связывания с БРЦ примыкает к промотору, взаимодействие БРЦ с ферментом приводит к увеличению скорости первичного связывания и скорости формирования закрытого комплекса. В промоторах II класса, где участок связывания с БРЦ расположен внутри промотора, взаимодействие БРЦ с ферментом приводит к возрастанию скорости перехода от закрытого комплекса к открытому.

12.22 Трансляция может регулироваться

Ключевые понятия:

- Белок-репрессор может регулировать трасляцию, препятствуя связыванию рибосомы с инициирующим кодоном.
- С помощью изменений в структуре полицистронной мРНК, возникающих в процессе трансляции, можно контролировать доступность инициирующих кодонов.

Трансляционный контроль свойственен оперонам, кодирующим компоненты аппарата белкового синтеза. В опероне происходят перестройки, обеспечивающие *координированную* регуляцию группы структурных генов. Дальнейший контроль также возложен на оперон. Однако уровень трансляции тоже может оказывать влияние на степень экспрессии отдельных генов.

В нескольких системах для трансляционного контроля используются схожие механизмы. *Репрессия трансляции происходит при связывании мРНК с белком, препятствующим узнаванию рибосомой инициирующей последовательности.* Это явление эквивалентно связыванию белка-репрессора с ДНК, в результате чего РНК-полимераза не может взаимодействовать с промотором. На РИС. 12.33 приведен наиболее общий механизм такого взаимодействия, в котором регуляторный белок связывает последовательность мРНК, содержащую инициирующий AUG кодон, экранируя ее от взаимодействия с рибосомой.

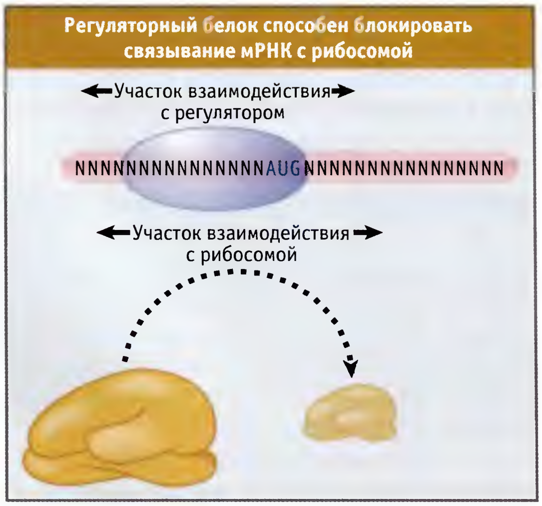


РИС. 12.33 Регуляторный белок способен взаимодействовать с участком мРНК, перекрывающимся с участком связывания рибосомы в области инициирующего кодона, и тем самым блокировать трансляцию

Некоторые примеры репрессоров трансляции и их мишеней приведены на **РИС. 12.34**. Классическим примером является поверхностный белок фага R17; он связывается со шпилькой, образованной последовательностью мРНК фага, которая включает в себя участок связывания с рибосомой. Похожим образом белок RegA фага T4 связывается с консенсусной последовательностью некоторых ранних мРНК фага, включающей инициирующий кодон AUG. T4 ДНК-полимераза взаимодействует с последовательностью своей собственной мРНК, содержащей необходимый для связывания рибосомы элемент Шайна–Дальгарно.

Другая форма трансляционного контроля наблюдается в том случае, когда трансляция одного цистрона требует изменений во вторичной структуре, зависящей от трансляции предыдущего цистрона. Это явление свойственно трансляции фагов с РНК-геномом, чьи цистроны всегда транслируются в определенном порядке. Как показано на РИС. 12.35, РНК фага формирует такую вторичную структуру, в которой доступной для узнавания остается только одна из инициирующих последовательностей; вторая не может быть узнана рибосомами, поскольку спарена с другим участком РНК. В ходе трансляции первого цистрона такая вторичная структура нарушается, и инициирующий сайт следующего цистрона становится доступным для взаимодействия с рибосомами. В данном случае способность мРНК транслироваться контролируется ее вторичной структурой.

Репрессоры трансляции связываются с мРНК

Репрессор	Ген-мишень	Сайт связывания
Поверхностный белок R17	R17 репликаза	Шпилька, включающая в себя участок связывания с рибосомой
T4 RegA	Ранние мРНК T4	Различные последовательности, в том числе и область инициирующего кодона
T4 ДНК-полимераза	T4 ДНК-полимераза	Последовательность Шайна–Дальгарно
T4 p32	Ген 32	Одноцепочечная 5′-лидирующая последовательность

РИС. 12.34 Белки, связывающиеся с последовательностями мРНК, расположенными в области, инициации трансляции, могут выступать в роли репрессоров трансляции

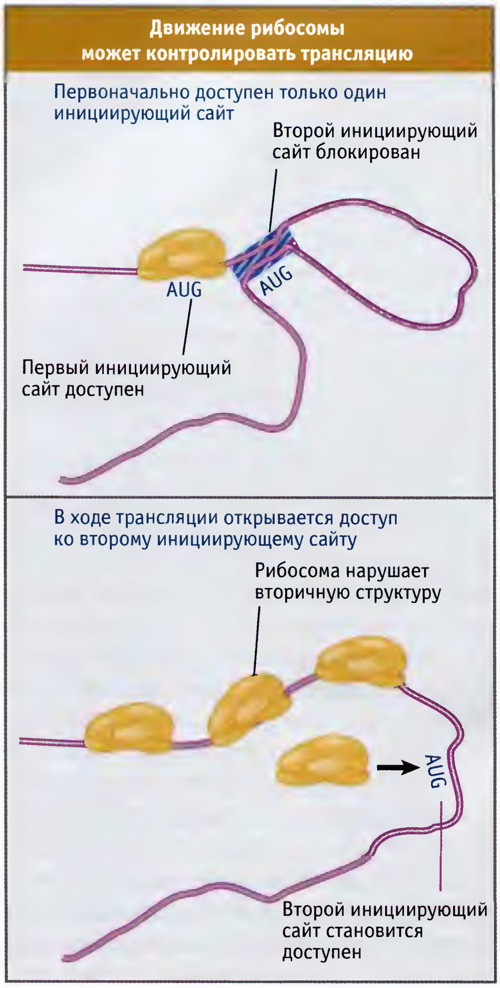


РИС. 12.35 **Реакция инициации может контролироваться вторичной структурой. Сначала (в результате определенной вторичной структуры молекулы мРНК) для узнавания рибосомой доступна только одна инициирующая последовательность. В ходе трансляции первого цистрона вторичная структура нарушается и для взаимодействия с рибосомами становится доступен другой инициирующий сайт (сайты)**

12.23 Трансляция рибосомальных белков контролируется аутогенно

Ключевые понятия:

- Трансляция рибосомальных белков контролируется одним из продуктов собственного оперона, связывающимся с полицистронной мРНК.

Аппарат, ответственный за экспрессию бактериальных генов, включает в себя около 70 белков. Главными его компонентами являются рибосомальные белки (наряду с дополнительными белками, участвующими в белковом синтезе). Остальная часть этого аппарата представлена субъединицами РНК-полимеразы и ее вспомогатльеными факторами. Все гены, кодирующие рибосомальные белки, факторы белкового синтеза и субъединицы РНК-полимеразы расположены вперемежку и организованы в небольшое число оперонов. Большинство этих белков кодируется только одним геном в хромосоме *E. coli.*

Координированный контроль способствует синтезу рибосомальных белков в количествах, соответствующих условиям роста: при большей скорости роста бактерии тратят больше усилий на формирование аппарата экспрессии генов. Для контроля экспрессии генов, кодирующих белки этого аппарата, а также для того, чтобы уровень синтеза рибосомальных белков соответствовал уровню синтеза рРНК, используется множество механизмов.

Данные об организации шести оперонов представлены на РИС. 12.36. Примерно половина генов рибосомальных белков (**r-белков**) картируется в четырех оперонах, тесно сцепленных друг с другом (их обозначают как *str, spc, S10* и α, каждый назван по первой из идентифицированных функций). Опероны *rif* и *L11* тоже сцеплены, но находятся в другой области генома.

Каждый оперон кодирует ряд функций. *str*-оперон несет гены рибосомальных белков малой субъединицы, а также факторов EF-Tu и EF-G. В двух оперонах, *spc* и *S10,* перемежаются гены белков как большой, так и малой рибосомальных субъединиц. В α-опероне находятся гены белков обеих субъединиц рибосомы и ген α-субъединицы РНК-полимеразы. В состав *rif*-оперона входят гены белков большой субъединицы рибосомы и гены β- и β′-субъединиц РНК-полимеразы.

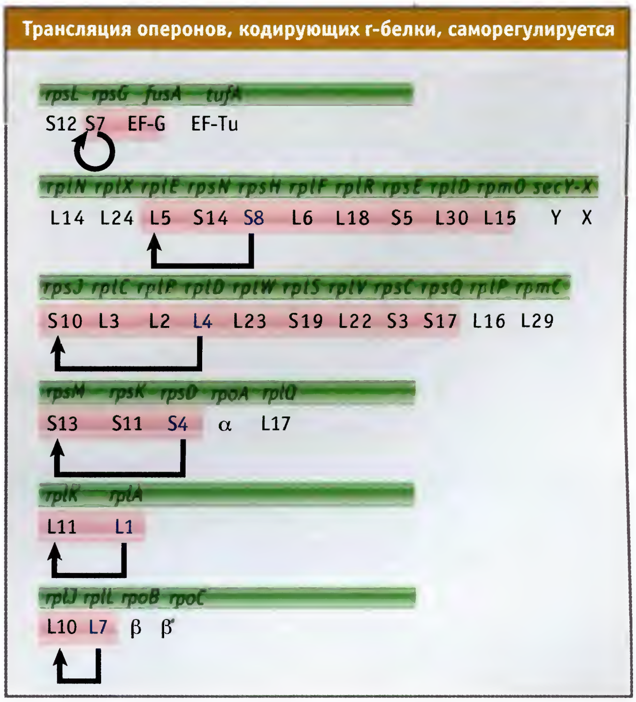


РИС. 12.36 **Гены рибосомальных белков, факторов белкового синтеза и субъединиц РНК-полимеразы организованы в небольшое число саморегулирующихся оперонов. Регуляторы обозначены синим цветом; регулируемые белки даны на розовом фоне**

Все рибосомальные белки, за исключением одного, требуются в эквимолярных количествах, что можно координировать уровнем рРНК. Разброс генов, продукты которых должны быть представлены в эквимолярных количествах, и их перемежаемость с генами, продукты которых необходимы в разных количествах, порождают определенные проблемы в реализации координированной регуляции.

Общим свойством всех оперонов, указанных на рис. 12.36, является регуляция некоторых генов одним из продуктов. Ген, кодирующий регуляторный белок, всегда сам является мишенью для своего продукта. **Аутогенной** называют тип такой регуляции, при которой белок (или РНК) регулирует свою собственную продукцию. Регуляторный белок ингибирует экспрессию группы генов, расположенных в пределах оперона, кодирующего образование рибосомальных белков. Такой тип регуляции получил название негативной аутогенной регуляции.

В каждом случае имеет место *накопление белка, приводящее к ингибированию синтеза как самого белка, так и продуктов некоторых других генов*. Эффект репрессии часто проявляется на уровне трансляции полицистронной мРНК. Все регуляторы являются рибосомальными белками, которые связываются непосредственно с рРНК. *Их влияние на трансляцию обусловлено также способностью связываться с собственной мРНК*. Сайты связывания регуляторных белков с мРНК либо перекрываются с последовательностью, где происходит инициация трансляции, либо расположены поблизости от нее. Возможно, вызывая конформационные изменения в структуре РНК, они влияют на доступность инициирующих кодонов для связывания. К примеру, белок L4 действует на точку старта трансляции в опероне S10, ингибируя трансляцию мРНК S10 и следующих за ним генов. Ингибирование может быть результатом простого блокирования доступа рибосомы, как это показано на рис. 12.34, либо эффект может сказываться на последующих стадиях трансляции. В двух случаях (описанных для оперонов S4 и α) регуляторный белок стабилизирует определенную вторичную структуру мРНК, которая делает невозможной продолжение реакции инициации после связывания мРНК с 30S субъединицей рибосомы.

Факт связывания r-белков с рРНК для установления аутогенной регуляции позволяет предположить механизм связи синтеза r-белков и рРНК. Обобщенная модель приведена на рис. 12.37. Допустим, что связь, образующаяся между регуляторным r-белком и рРНК, гораздо более прочная, чем связь между r-белком и мРНК. В этом случае новосинтезируемые r-белки будут связываться с рРНК до тех пор, пока будет иметься свободная рРНК, что, в свою очередь, будет приводить к сборке рибосом. Свободного r-белка, способного связаться с мРНК, не будет, в результате трансляция будет продолжаться. Однако как только синтез рРНК замедлится или остановится, свободные r-белки начнут накапливаться. Они смогут связываться с мРНК и репрессировать дальнейшую трансляцию. Такой цикл гарантирует одинаковый ответ каждого оперона, детерминирующего синтез r-белка, на содержание рРНК. В случае избытка r-белка над рРНК, синтез белка будет подавлен.

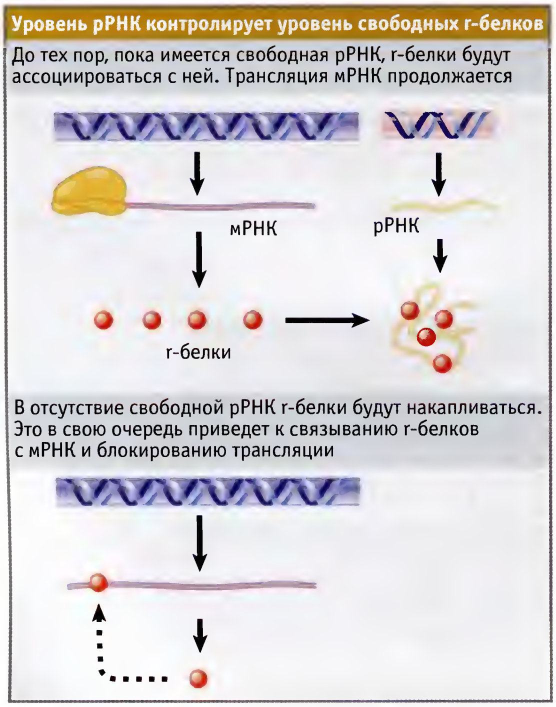


Рис. 12.37 Трансляция оперонов, кодирующих r-белки, регулируется аутогенно и зависит от уровня рРНК

12.24 Синтез белка p32 фага T4 контролируется аутогенно

Ключевые понятия:

- p32 связывается со своей собственной мРНК, блокируя инициацию трансляции.

Первым примером аутогенной регуляции явились данные, полученные при изучении гена, кодирующего белок p32 фага T4. Этот белок играет центральную роль в процессах генетической рекомбинации, репарации и репликации ДНК, в которых его функция реализуется за счет его способности связываться с одноцепочечной ДНК. Нонсенс-мутации в гене *32* ведут к перепроизводству неактивного белка. *Это означает, что в тех случаях, когда функция белка нарушена, он синтезируется в больших количествах*. Данный эффект проявляется на уровне трансляции; мРНК гена *32* стабильна и сохраняется независимо от поведения белкового продукта.

На рис. 12.38 представлена модель, согласно которой осуществляется контроль работы гена *32*. Если в

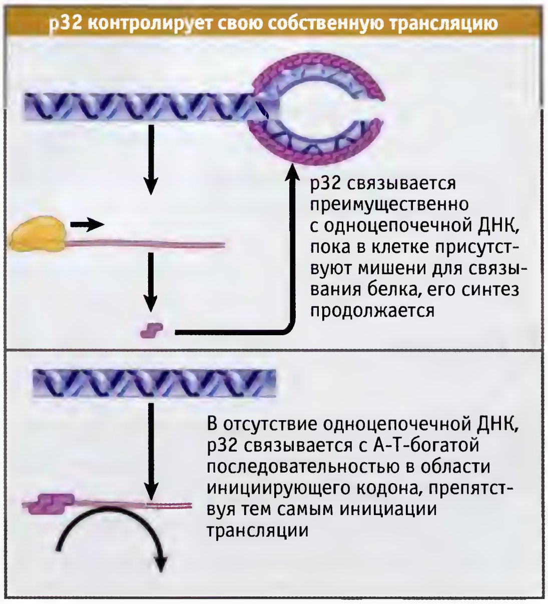


РИС. 12.38 **Избыток белкового продукта гена *32* (p32) связывается со своей собственной мРНК и блокирует инициацию трансляции рибосомой**

инфицированной фагом клетке присутствует одноцепочечная ДНК, она будет связывать белок p32. Однако в отсутствие одноцепочечной ДНК или, по крайней мере, в условиях избытка p32 белок предотвращает трансляцию своей собственной мРНК. Репрессия реализуется в результате прямого связывания p32 с мРНК, что мешает инициации трансляции. По всей вероятности, это происходит в А-Т-богатом регионе, окружающем сайт связывания с рибосомой на мРНК.

Две особенности процесса связывания p32 с мРНК влияют на эффективность системы контроля:



РИС. 12.39 **Белковый продукт гена *32* связывается с разными субстратами с разной аффинностью. Согласно прочности связывания с белком, эти субстраты можно расположить в следующем порядке: одноцепочечная ДНК, мРНК p32, случайная мРНК. Связывание со своей собственной мРНК не дает концентрациям белка превышать 10^{-6} М**

- Аффинность p32 к мРНК должна быть существенно ниже, чем к одноцепочечной ДНК. Константа равновесия для реакции связывания с РНК почти на два порядка ниже величины этого же параметра для реакции связывания с одноцепочечной ДНК.
- Аффинность p32 к мРНК, однако, должна быть существенно выше его аффинности к другим РНК-последовательностям. Это зависит от нуклеотидного состава и вторичной структуры РНК; важной особенностью региона мРНК, соответствующего гену *32*, является отсутствие вторичной структуры на довольно протяженном участке.

С помощью известных констант равновесия мы можем представить связывание p32 с различными мишенями как функцию, зависящую от концентрации белка. Как показано на РИС. 12.39, при концентрациях p32 ниже 10^{-6} М он связывается с одноцепочечной ДНК. При концентрациях $>10^{-6}$ М p32 связывается с мРНК. При дальнейшем увеличении концентрации белок связывается с другими РНК с различной степенью аффинности.

На основе этих данных можно сказать, что уровень p32 должен саморегулироваться таким образом, чтобы не превышать 10^{-6} М, что соответствует ~ 2000 молекулам на бактериальную клетку. Этому условию удовлетворяет уровень белка от 1000 до 2000 молекул на клетку.

Особенность аутогенного контроля заключается в том, что каждое регуляторное взаимодействие уникально: белок действует только на свою собственную мРНК. Другим примером аутогенной регуляции является еще один, более общий, регулятор трансляции фага T4. Белок, кодируемый геном *regA*, угнетает экспрессию нескольких генов, транскрибируемых на ранних стадиях инфекции. RegA блокирует трансляцию этих генов, конкурируя с 30S субъединицей рибосомы за связывание с мРНК. Действие RegA является прямым аналогом функции репрессорного белка, способного связываться с несколькими операторами.

12.25 Аутогенный контроль и сборка макромолекул

Ключевые понятия:

- **Предшественник белка микротрубочек, свободный тубулин, ингибирует трансляцию тубулиновой мРНК.**

Аутогенный контроль — это общий тип контроля среди белков, участвующих в сборке макромолекулярных ансамблей. Сами собираемые частицы в качестве регулятора могут быть непригодны, так как они слишком крупные, слишком малочисленны, или имеют ограниченную локализацию. Однако потребность в синтезе компонентов комплекса может быть отражена пулом свободных субъединиц-предшественников. Если процесс сборки по

какой-то причине блокирован, то в результате накопления субъединиц дальнейший синтез компонентов, ставший теперь ненужным, прекращается.

Эукариотические клетки обладают системой, которой свойственна аутогенная регуляция такого типа. Тубулин представляет собой мономер, из которого собираются микротрубочки — главный компонент системы цитоскелета всех эукариотических клеток. Образование мРНК тубулина контролируется пулом свободного тубулина. При достижении определенной концентрации белка дальнейшее образование тубулиновой мРНК прекращается. Снова используется тот же принцип: тубулин, задействованный в процессе сборки макромолекул, не играет роли в регуляции, однако количество свободного предшественника определяет, будет ли продолжаться его синтез.

Мишенью для действия регуляторов является короткая нуклеотидная последовательность в начале кодирующего региона. Мы пока не знаем точной роли этой последовательности. На РИС. 12.40 приведены две возможные модели. Тубулин может связываться непосредственно с мРНК или с растущей полипептидной цепью. Независимо от применяемой модели, избыток тубулина приведет к деградации расположенной на полисомах тубулиновой мРНК. Следствием данной реакции становится потеря стабильности мРНК.

Система аутогенной регуляции является системой внутреннего самоограничения, в отличие от систем внешнего контроля, обсуждаемых нами ранее. Способность репрессорного белка связываться с оператором может регулироваться поступающими извне малыми молекулами, активирующими или ингибирующими репрессор. При аутогенной регуляции критическим параметром становится концентрация самого регуляторного белка.

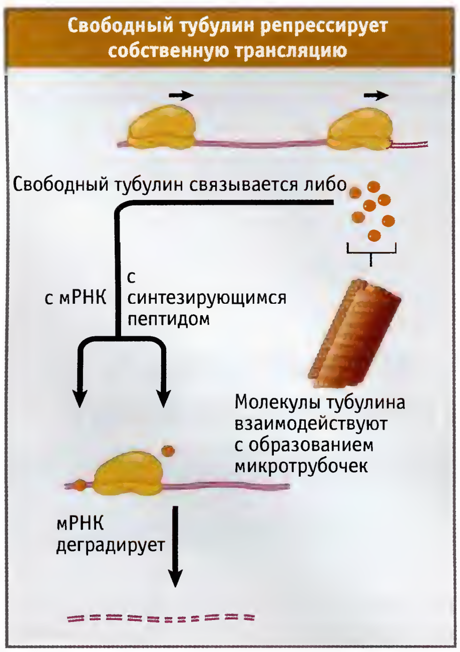


РИС. 12.40 **Тубулин представляет собой мономер, из которого собираются микротрубочки. Накопление избытка тубулина приводит к связыванию его либо с собственной мРНК в области начала рамки считывания, либо с синтезирующимся пептидом, что, в свою очередь, ведет к нестабильности тубулиновой мРНК**

12.26 Резюме

Регуляция транскрипции осуществляется за счет взаимодействия *транс*-активных факторов и *цис*-активных сайтов. В роли *транс*-активных факторов выступают продукты регуляторных генов. Как правило, они представляют собой белки, в некоторых случаях — РНК. Диффундируя в клетке, *транс*-активные факторы способны оказывать свой эффект на ген-мишень любой локализации. *Цис*-активные сайты на ДНК (или РНК) — это последовательности, основная функция которых — быть узнанными *in situ* регуляторными факторами. Они не кодируют белки и способны регулировать только расположенные рядом гены. Бактериальные гены, кодирующие белки с родственной функцией (как, например, последовательные ферменты одного пути метаболизма), как правило, организованы в кластеры и транскрибируются с одного промотора в полицистронную мРНК. Контролируя работу одного промотора, можно регулировать экспрессию всех генов кластера. Регуляторную единицу, содержащую в своем составе структурные гены и *цис*-активные элементы, называют опероном.

Инициация транскрипции регулируется в результате серии событий, происходящих в области промотора. Способность РНК-полимеразы инициировать транскрипцию на промоторе блокируется или усиливается с помощью других белков. Гены, активные до тех пор, пока не будут выключены с помощью регуляторных факторов, подвержены негативному контролю. Гены, экспрессия которых включается только под действием регуляторных факторов, подвержены позитивному контролю. Тип контроля может быть определен путем анализа характера взаимодействия между аллелями регуляторных генов дикого типа и мутантными аллелями, которые могут приводить к конститутивной экспрессии структурных генов (дерепрессия, гены постоянно включены) или к их неиндуцибельности (суперрепрессия, гены постоянно выключены).

Белок-репрессор препятствует либо связыванию РНК-полимеразы с промотором, либо активации транскрипции. Репрессор связывается с последовательностью-мишенью (оператором), которая расположена, как правило, в районе или выше стартовой точки транскрипции. Оператор представляет собой короткую (и часто палиндромную) последовательность. Репрессор — гомомультимер, чья симметрия отражает особенности структуры последовательности-мишени.

Способность репрессора связываться с оператором контролируется малыми молекулами. Индуктор препятствует связыванию репрессора с оператором; корепрессор — способствует. Взаимодействие индуктора и корепрессора с репрессором приводит к изменению структуры его ДНК-связывающего сайта. Такой аллостерический контроль возможен как в свободных молекулах репрессора, так и в уже связанных с ДНК.

Метаболизм лактозы запускается в результате индукции. Индуктор — β-галактозид, связываясь с репрессором, мешает его взаимодействию с *lac*-оператором, в результате чего запускается транскрипция и трансляция гена *lacZ*, кодирующего β-галактозидазу — фермент, метаболизирующий β-галактозиды. Метаболизм триптофана контролируется с помощью репрессии. Корепрессор (триптофан) активирует репрессорный белок, в результате чего тот связывается с *trp*-оператором и блокирует экспрессию генов, кодирующих ферменты биосинтеза триптофана. Репрессор может оказывать свое влияние сразу на несколько мишеней, если они содержат в своем составе копию консенсусной последовательности оператора.

Белок, обладающий высокой аффинностью к определенным последовательностям ДНК, обладает и слабым сродством ко всем остальным последовательностям ДНК. Отношение этих двух показателей определяет специфичность белка. Неспецифичных сайтов связывания (случайных последовательностей ДНК) в геноме значительно больше, чем специфичных. В связи с этим ДНК-связывающие белки, такие как репрессор или РНК полимераза, «хранятся» ассоциированными с неспецифичными последовательностями ДНК. (По всей видимости, свободных молекул таких белков в клетке почти нет.) Специфичность белка по отношению к его мишени должна быть очень высока для того, чтобы компенсировать избыток случайных сайтов. Количество белков в бактериальной клетке и их специфичность отрегулированы таким образом, чтобы ген во включенном состоянии обеспечивал узнавание и связывание со специфичной мишенью, а в выключенном — полное освобождение мишени от регуляторного фактора.

Литература

12.1 Введение

Исследование

Jacob, F. and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* 3, 318–389.

12.2 Негативная и позитивная регуляция экспрессии генов

Обзор

Miller, J. and Reznikoff, W., eds. (1980). *The Operon*, 2nd ed., Woodbury, NY: Cold Spring Harbor Laboratory Press.

12.4 Гены *lac* контролируются репрессором

Обзор

Barkley, M. D. and Bourgeois, S. (1978). Repressor recognition of operator and effectors. In *The Operon*, eds. Miller, J. and Reznikoff, W. New York: Cold Spring Harbor Laboratory, 177–220.

Beckwith, J. (1978). lac: the genetic system. In *The Operon*, eds. Miller, J. and Reznikoff, W. New York: Cold Spring Harbor Laboratory, 11–30.

Beyreuther, K. (1978). Chemical structure and functional organization of the lac repressor from *E. coli.* In *The Operon*, eds. Miller, J. and Reznikoff, W. New York: Cold Spring Harbor Laboratory, 123–154.

Miller, J. H. (1978). The *lacl* gene: its role in lac operon control and its use as a genetic system. In *The Operon*, eds. Miller, J. and Reznikoff, W. New York: Cold Spring Harbor Laboratory, 31–88.

Weber, K. and Geisler, N. (1978). Lac repressor fragments produced in vivo and in vitro: an approach to the understanding of the interaction of repressor and DNA. In *The Operon*, eds. Miller, J. and Reznikoff, W. New York: Cold Spring Harbor Laboratory, 155–176.

Wilson, C. J., Zahn, H., Swint-Kruse, L., and Matthews, K. S. (2006). The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding. *Cell. Mol. Life Sci.* November 13.

Исследование

Jacob, F. and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* 3, 318–389.

12.10 Субъединица репрессора состоит из нескольких доменов

Исследование

Friedman, A. M., Fischmann, T. O., and Steitz, T. A. (1995). Crystal structure of lac repressor core tetramer and its implications for DNA looping. *Science* 268, 1721–1727.

Lewis, M. et al. (1996). Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* 271, 1247–1254.

12.13 Мутантные фенотипы коррелируют со структурой отдельных доменов репрессора

Обзор

Pace, H. C., Kercher, M. A., Lu, P., Markiewicz, P., Miller, J. H., Chang, G., and Lewis, M. (1997). Lac repressor genetic map in real space. *Trends Biochem. Sci.* 22, 334–339.

Markiewicz, P., Kleina, L. G., Cruz, C., Ehret, S., and Miller, J. H. (1994). Genetic studies of the lac repressor. XIV. Analysis of 4000 altered *E. coli lac* repressors reveals essential and non-essential residues, as well as spacers which do not require a specific sequence. *J. Mol. Biol.* 240, 421–433.

Suckow, J., Markiewicz, P., Kleina, L. G., Miller, J., Kisters-Woike, B., and Müller-Hill, B. (1996). Genetic studies of the Lac repressor. XV: 4000 single amino acid substitutions and analysis of the resulting phenotypes on the basis of the protein structure. *J. Mol. Biol.* 261, 509–523.

12.14 Репрессорный белок связывается с оператором

Исследование

Gilbert, W. and Müller-Hill, B. (1966). Isolation of the lac repressor. *Proc. Natl. Acad. Sci.* USA 56, 1891–1898.

Gilbert, W. and Müller-Hill, B. (1967). The lac operator is DNA. *Proc. Natl. Acad. Sci. USA* 58, 2415–2421.

12.16 Репрессор способен связывать три оператора и взаимодействовать с РНК-полимеразой

Исследование

Oehler, S. et al. (1990). The three operators of the lac operon cooperate in repression. *EMBO J.* 9, 973–979.

12.18 Оператор конкурирует со случайными последовательностями за связывание с репрессором

Исследование

Cronin, C. A., Gluba, W., and Scrable, H. (2001). The lac operator-repressor system is functional in the mouse. *Genes Dev.* 15, 1506–1517.

Hildebrandt, E. R. et al. (1995). Comparison of recombination *in vitro* and in *E. coli* cells: measure of the effective concentration of DNA *in vivo*. *Cell* 81, 331–340.

Lin, S.-Y. and Riggs, A. D. (1975). The general affinity of lac repressor for *E. coli* DNA: implications for gene regulation in prokaryotes and eukaryotes. *Cell* 4, 107–111.

12.21 БРЦ функционирует на основе различных механизмов, действуя при этом на разные опероны

Обзор

Botsford, J. L. and Harman, J. G. (1992). Cyclic AMP in prokaryotes. *Microbiol. Rev.* 56, 100–122.

Kolb, A. (1993). Transcriptional regulation by cAMP and its receptor protein. *Annu. Rev. Biochem.* 62, 749–795.

Исследование

Niu, W., Kim, Y., Tau, G., Heyduk, T., and Ebright, R. H. (1996). Transcription activation at class II CAP-dependent promoters: two interactions between CAP and RNA polymerase. *Cell* 87, 1123–1134.

Zhou, Y., Busby, S., and Ebright, R. H. (1993). Identification of the functional subunit of a dimeric transcription activator protein by use of oriented heterodimers. *Cell* 73, 375–379.

Zhou, Y., Merkel, T. J., and Ebright, R. H. (1994). Characterization of the activating region of *E. coli* catabolite gene activator protein (CAP). II. Role at Class I and class II CAP-dependent promoters. *J. Mol. Biol.* 243, 603–610.

12.23 Трансляция рибосомальных белков контролируется аутогенно

Обзор

Nomura, M. et al. (1984). Regulation of the synthesis of ribosomes and ribosomal components. *Annu. Rev. Biochem.* 53, 75–117.

Исследование

Baughman, G. and Nomura, M. (1983). Localization of the target site for translational regulation of the L11 operon and direct evidence for translational coupling in *E. coli*. *Cell* 34, 979–988.

12.25 Аутогенный контроль и сборка макромолекул

Обзор

Gold, L. (1988). Posttranscriptional regulatory mechanisms in *E. coli*. *Annu. Rev. Biochem.* 57, 199–223.

Регуляторная РНК

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

13.1 Введение

- Формируя участки определенной вторичной структуры, РНК реализует свою регуляторную функцию. При взаимодействии с последовательностью-мишенью эти участки изменяют ее свойства. Вторичная структура может быть следствием как внутри-, так и межмолекулярных взаимодействий.

13.2 Аттенуация контролируется с помощью альтернативных вторичных структур

- Терминация транскрипции может быть аттенуирована путем формирования шпильки во вторичной структуре РНК.
- В механизмах аттенуации задействованы белки, стабилизирующие или дестабилизирующие шпильку во вторичной структуре РНК.

13.3 Терминация *trp*-генов у *Bacillus subtilis* контролируется уровнями триптофана и тРНК^{Trp}

- Терминаторный белок, называемый TRAP, активируется триптофаном и блокирует транскрипцию генов *trp*.
- Активность TRAP (опосредованно) ингибируется тРНК^{Trp}, не несущей аминокислотного остатка.

13.4 Триптофановый оперон *E. coli* контролируется с помощью аттенуации

- Аттенуатор (ρ-независимый терминатор) располагается между промотором и первым геном кластера *trp*.
- В отсутствие триптофана терминация супрессируется, а эффективность транскрипции возрастает в 10 раз.

13.5 Аттенуация контролируется с помощью трансляции

- Лидерная последовательность *trp*-оперона представлена открытой рамкой считывания, состоящей из 14 кодонов, два из которых являются кодонами триптофана.
- Структура РНК в аттенуаторе зависит от того, транслируется ли эта рамка считывания.
- В присутствии триптофана лидерная последовательность транслируется в белок, в области аттенуатора формируется шпилька, и транскрипция терминируется.
- В отсутствие триптофана рибосома останавливает трансляцию, достигнув триптофановых кодонов, а формирующаяся при этом альтернативная вторичная структура мРНК препятствует образованию шпильки и транскрипция продолжается.

13.6 Инактивация экспрессии генов с помощью антисмысловой РНК

- Антисмысловые гены при введении их в эукариотические клетки блокируют экспрессию генов-мишеней.

13.7 Регуляция трансляции с помощью молекул малых РНК

- Регуляторная РНК реализует свои функции, формируя дуплекс с РНК-мишенью.
- Дуплекс может блокировать инициацию транскрипции, вызывать терминацию транскрипции, или служить мишенью для действия эндонуклеаз.

13.8 Регуляторные РНК бактерий

- Бактериальные регуляторные РНК называют малыми РНК.
- Некоторые из бактериальных малых РНК находятся в связанном с белком Hfq состоянии, что повышает их эффективность.
- OxyS sРНК на посттрансляционном уровне активирует или репрессирует экспрессию более чем 10 локусов.

13.9 МикроРНК являются регуляторами во многих эукариотических организмах

- Геномы животных и растений кодируют большое количество коротких (~22 нуклеотида) молекул РНК, названных микроРНК.
- МикроРНК регулируют экспрессию генов, связываясь с комплементарными последовательностями мРНК-мишеней.

13.10 РНК-интерференция — причина молчания генов

- В результате РНК-интерференции запускается процесс деградации мРНК, комплементарной любой из цепей короткой двухцепочечной регуляторной РНК (дцРНК).
- Двухцепочечная РНК способна подавлять экспрессию генов, приводя к их молчанию.

13.11 Резюме

13.1 Введение

Ключевые понятия:

- Формируя участки определенной вторичной структуры, РНК реализует свою регуляторную функцию. При взаимодействии с последовательностью-мишенью эти участки изменяют ее свойства. Вторичная структура может быть следствием как внутри-, так и межмолекулярных взаимодействий.

Контроль экспрессии генов осуществляется при взаимодействии регуляторных факторов со специфическими последовательностями ДНК или мРНК на стадиях, предшествующих синтезу белка. В этом заключается основной принцип механизмов контроля у бактерий. Регулируемой стадией может быть транскрипция (при этом мишенью для действия регулятора будет служить последовательность ДНК) или трансляция (в этом случае в роли мишени будет выступать РНК). Контроль транскрипции может осуществляться на двух ее стадиях — инициации и терминации. Регуляторным фактором может быть белок или РНК. Регулятор способен выключать (репрессировать) либо включать (активировать) экспрессию последовательности-мишени. Экспрессия многих генов координированно регулируется одним регуляторным геном. Такой механизм контроля предполагает, что все контролируемые гены имеют какую-то одинаковую последовательность или структуру, распознаваемую регулятором. Экспрессия самих регуляторов тоже контролируется. Как правило, ее уровень меняется в зависимости от присутствия малых молекул, состав которых отражает условия окружающей среды. Кроме того, регуляторы могут находиться под контролем других регуляторов, формируя таким образом сложные регуляторные сети.

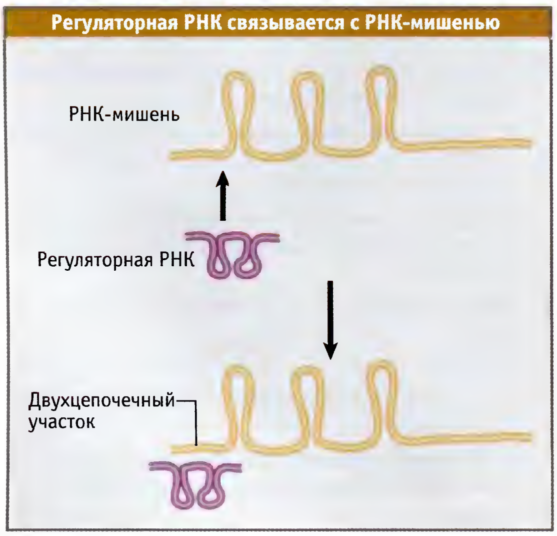


РИС. 13.1 Регулятор — малая молекула РНК, содержащая в своем составе одноцепочечный участок, комплементарный одноцепочечному участку на РНК-мишени

Попытаемся сравнить механизм работы регуляторов разных типов.

Действие белковых регуляторов основано на принципе аллостерической регуляции. Белок имеет два сайта связывания: один — для нуклеотидной последовательности-мишени, другой — для малой молекулы. Связывание малой молекулы в предназначенном для этого участке изменяет конформацию регулятора, а это, в свою очередь, влияет на аффинность связывания нуклеиновой кислоты в другом сайте. Данный процесс детально описан для репрессора Lac. Белковые регуляторы часто имеют мультимерную структуру с симметричной организацией, что позволяет двум субъединицам одновременно контактировать с палиндромной последовательностью-мишенью на ДНК. Суммарный эффект связывания повышает чувствительность к регуляторным влияниям.

Главным принципом РНК-опосредованной регуляции является изменение вторичной структуры. Способность РНК менять свои регуляторные свойства, меняя конформацию, является альтернативой аллостерической регуляции активности белков. Изменения в структуре могут быть следствием как меж-, так и внутримолекулярных взаимодействий.

Внутримолекулярные взаимодействия, приводящие к появлению новой вторичной структуры, представляют собой различные варианты спаривания оснований в молекуле РНК. Свойства альтернативных конформаций могут отличаться. Изменения во вторичной структуре молекулы мРНК могут влиять на ее способность транслироваться. Для осуществления терминации транскрипции необходимо участие элементов вторичной структуры мРНК, поэтому конформационные изменения молекулы мРНК могут сказываться на процессе терминации.

При межмолекулярных взаимодействиях РНК-регулятор распознает свою мишень по тому же принципу комплементарного спаривания оснований. Как показано на РИС. 13.1, регулятором обычно является малая молекула РНК с определенной вторичной структурой и одноцепочечным участком, комплементарным одноцепочечному участку РНК-мишени. Формирование дуплекса между регуляторной РНК и РНК-мишенью приводит к следующим последствиям:

- Формирование двуспиральной структуры само по себе может быть достаточным регуляторным событием. В одних случаях белок (белки) могут взаимодействовать только с одноцепочечной РНК, поэтому формирование дуплекса блокирует их функцию. В других случаях, наоборот, мишенью для связывания служит участок двухцепочечной структуры; способностью связываться с такими регионами обладают, например, нуклеазы, осуществляющие деградацию мРНК и препятствующие ее трансляции.
- Дуплекс между мишенью и регулятором может иметь значение при формировании альтернативной вторичной структуры молекулы.

13.2 Аттенуация контролируется с помощью альтернативных вторичных структур

Ключевые понятия:

- Терминация транскрипции может быть аттенуирована путем формирования шпильки во вторичной структуре РНК.
- В механизмах аттенуации задействованы белки, стабилизирующие или дестабилизирующие шпильку во вторичной структуре РНК.

Регуляция, опосредованная изменением вторичной структуры РНК, происходит как у эукариот, так и прокариот. Реализуя различные варианты спаривания оснований в молекуле, РНК может формировать альтернативные вторичные структуры. Свойства альтернативных конформаций могут отличаться. Такой тип регуляции характерен для контроля терминации транскрипции, поскольку разные вторичные структуры мРНК отличаются по своей способности вызывать терминацию. Другим способом контроля конформационного состояния (а значит и функции) является расщепление РНК; при отщеплении отдельных сегментов РНК конформация остальных может меняться. Осуществление контроля активности РНК-мишени возможно при спаривании по принципу комплементарности с малой регуляторной молекулой РНК; роль малых молекул РНК в этом случае аналогична роли регуляторных белков (см. разд. 13.7 «Регуляция трансляции с помощью малых молекул РНК»). Способность РНК, изменяя конформацию, менять и свои регуляторные свойства является альтернативой аллостерическому контролю активности регуляторных белков. Взаимодействия в одном участке молекулы влияют на структуру других регионов с помощью двух этих механизмов.

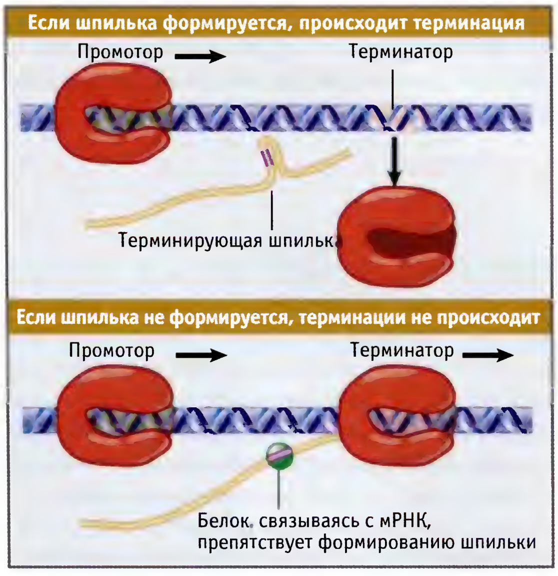


РИС. 13.2 Аттенуация наблюдается в том случае, когда терминирующая шпилька на РНК по каким-то причинам не образуется

Некоторые опероны регулируются с помощью **аттенуации** — явления, в ходе которого реализуется способность РНК полимеразы прочитывать сквозь **аттенуаторы**. Аттенуаторы представляют собой истинные терминаторные последовательности, расположенные в начале транскрипционной единицы. *Принцип аттенуации заключается в контроле формирования шпильки, необходимой для осуществления ρ-независимой терминации, с помощью внешних факторов.* При образовании в структуре мРНК шпильки РНК-полимераза не может осуществлять транскрипцию структурных генов. В отсутствие шпильки полимераза продолжает транскрипцию за пределы терминатора, и гены экспрессируются. В разных системах контроля структуры РНК используются разные механизмы.

Аттенуация может контролироваться белками, связывающимися с РНК и стабилизирующими или дестабилизирующими ее вторичную структуру. На РИС. 13.2 показана ситуация, в которой белок препятствует формированию терминирующей шпильки. Белок может сам по себе обладать такой активностью, а может зависеть от малых регуляторных молекул тем же образом, каким репрессорный белок зависит от корепрессора.

13.3 Терминация *trp*-генов у *Bacillus subtilis* контролируется триптофаном и тРНК^{Trp}

Ключевые понятия:

- Терминаторный белок, называемый TRAP, активируется триптофаном и блокирует транскрипцию генов *trp*.
- Активность TRAP (опосредованно) ингибируется тРНК^{Trp}, не несущей аминокислотного остатка.

Регуляторная сеть, контролирующая транскрипцию на стадии терминации, может задействовать механизмы как прямого, так и косвенного влияния малых молекул продуктов или субстратов.

У *Bacillus subtilis* белок, названный аттенуирующим белком, связывающим мРНК триптофана (TRAP, ранее MtrB), активируясь триптофаном (Trp), связывается с лидирующей последовательностью синтезируемого транскрипта. TRAP формирует мультимер из двенадцати субъединиц. Каждая субъединица связывает одну аминокислоту (триптофан) и триплетную последовательность на мРНК (GAG или UAG). РНК в месте связывания с белком сворачивается в петлю. Как показано на РИС. 13.3, этот процесс гарантирует доступность участков мРНК, участвующих в формировании терминирующей шпильки. В результате терминации транскрипции прекращается продукция фермента, ответственного за синтез триптофана. В действительности TRAP служит терминаторным белком, активность которого зависит от уровня триптофана. В отсутствие TRAP мРНК сворачивается в альтернативную вторичную структуру, препятствующую формированию терминирующей шпильки.

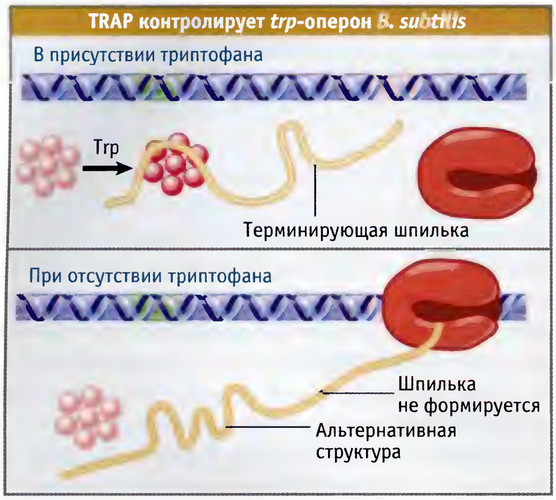


РИС. 13.3 Активируясь триптофаном, TRAP связывает мРНК. При этом формируется терминирующая шпилька, с помощью которой РНК-полимераза терминирует транскрипцию (и гены не экспрессируются). В отсутствие триптофана TRAP не связывается с мРНК, и мРНК приобретает альтернативную вторичную структуру, которая препятствует формированию терминирующей шпильки

TRAP, в свою очередь, контролируется тРНК^{Trp}. На РИС. 13.4 показано, как тРНК^{Trp} связывается с мРНК белка, названного анти-TRAP (AT). Связываясь с TRAP, этот белок мешает формированию терминирующей шпильки на мРНК. Свободная тРНК^{Trp} также повышает эффективность трансляции *trp*-мРНК. Суммарным результатом этих двух процессов является увеличение уровня синтеза белка анти-TRAP, который, связываясь с TRAP, не дает ему репрессировать триптофановый оперон. В отсутствие триптофана образуется избыток свободной тРНК^{Trp}, что способствует синтезу анти-TRAP, который, в свою очередь, блокируя функцию TRAP, приводит к экспрессии триптофановых генов.

Экспрессия *trp* генов у *Bacillus subtilis* контролируется как уровнем триптофана, так и уровнем тРНК^{Trp}. Если триптофан присутствует в клетке и синтезировать его не нужно, активируется TRAP и экспрессия ферментов синтеза триптофана ингибируется. Наличие свободной тРНК^{Trp} свидетельствует о недостатке триптофана. Свободная тРНК^{Trp} активирует анти-TRAP и способствует тем самым транскрипции генов *trp*.

13.4 Триптофановый оперон *E. coli* контролируется с помощью аттенуации

Ключевые понятия:

- Аттенуатор (независимый терминатор) располагается между промотором и первым геном кластера *trp*.
- В отсутствие триптофана терминация супрессируется, а эффективность транскрипции возрастает в 10 раз.

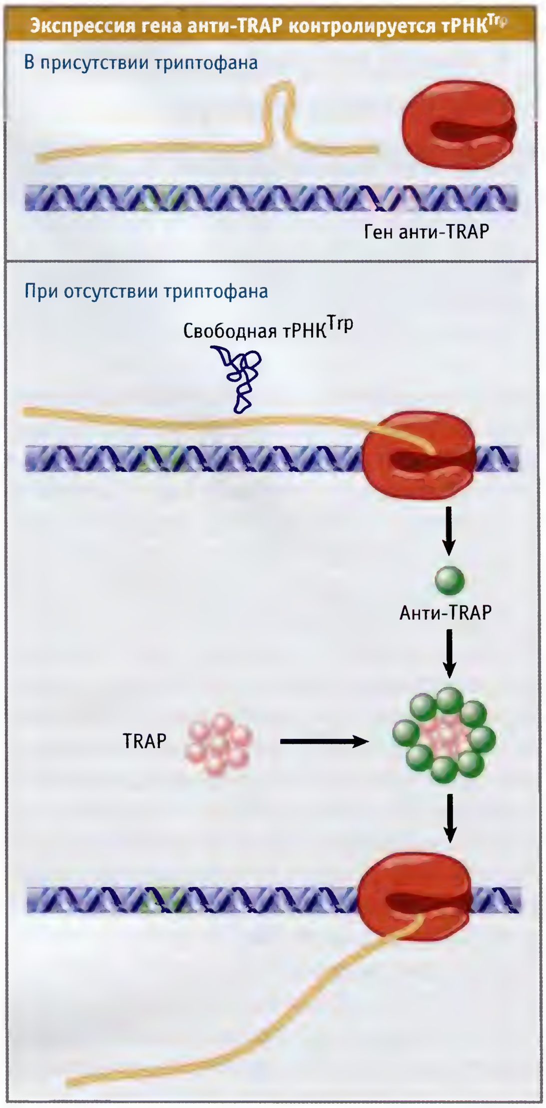


РИС. 13.4 В нормальных условиях (в присутствии триптофана) терминация транскрипции происходит перед геном анти-TRAP. В отсутствие триптофана, свободная тРНК^{Trp} связывается с анти-TRAP-мРНК, блокируя формирование терминирующей шпильки и способствуя экспрессии анти-TRAP

Для контроля работы генов в клетках *E. coli* (где впервые было обнаружено явление аттенуации) используется сложная регуляторная система. Изменения вторичной структуры мРНК зависят от положения рибосомы на мРНК. Как показано на РИС. 13.5, для осуществления терминации необходимо, чтобы *рибосома транслировала участок мРНК, расположенный перед генами* trp *(соответствующий лидирующей последовательности пептида)*. Когда рибосома транслирует лидирующий регион, в области терминатора 1 формируется терминирующая шпилька. В том случае, когда рибосома по каким-то причинам не может транслировать лидирующую последовательность, шпилька не образуется, терминации не происходит и

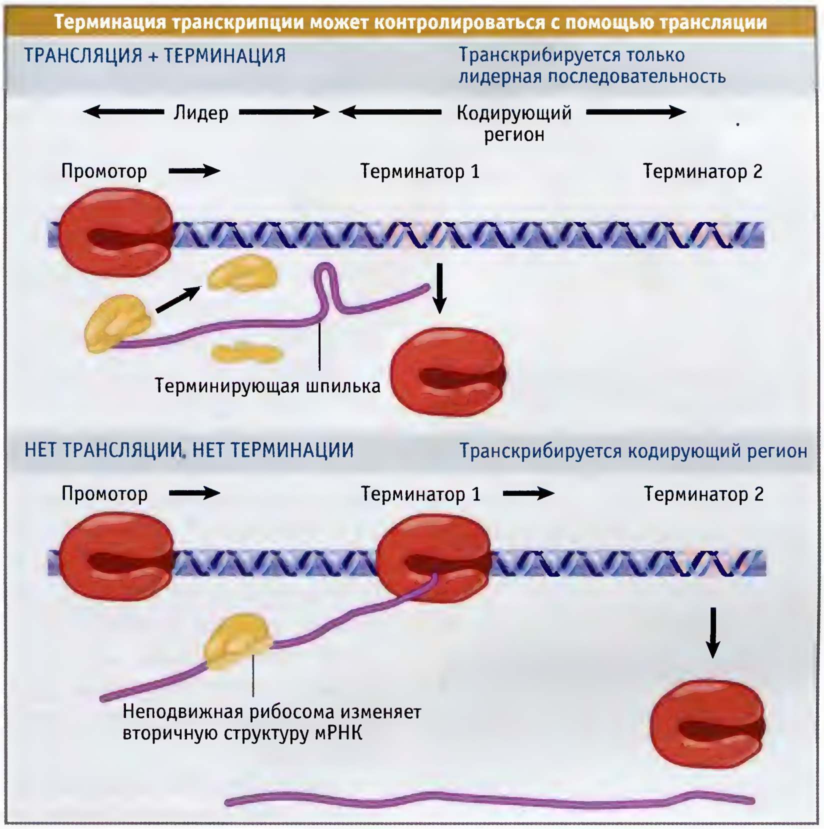


РИС. Вторичная структура РНК, влияя на движение вдоль мРНК рибосом, участвует тем самым в регуляции терминации транскрипции

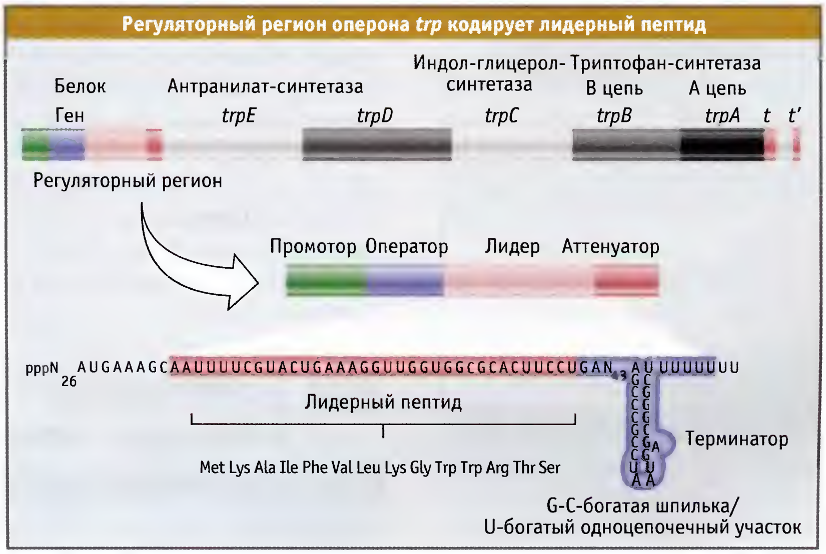
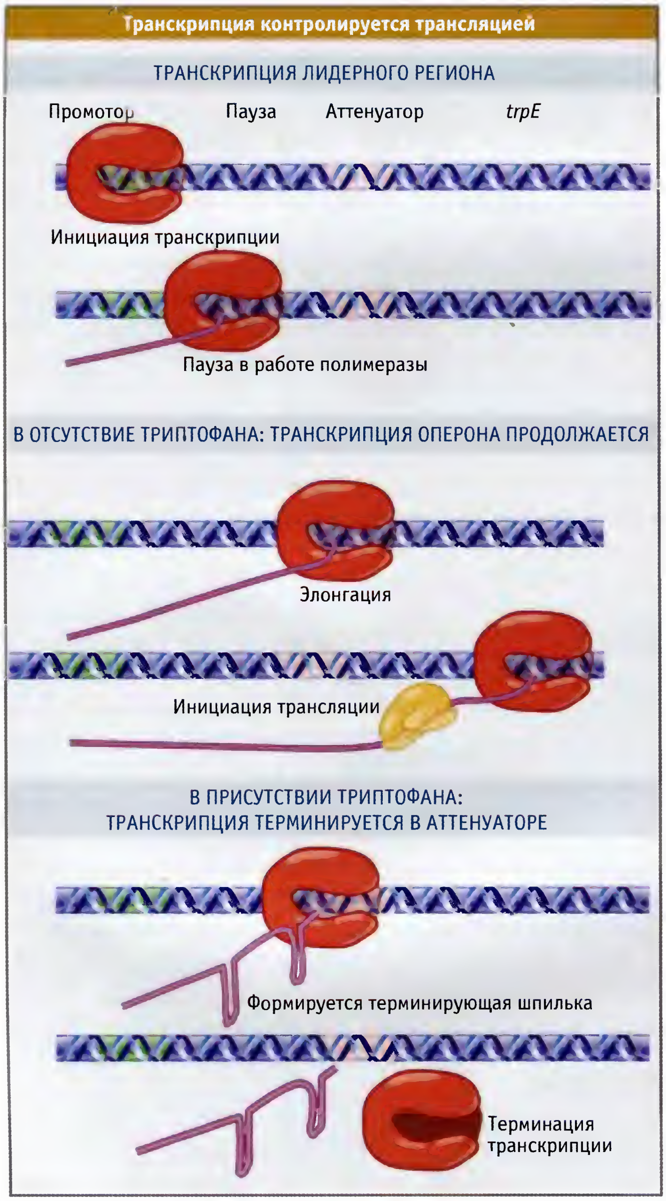


РИС. Оперон *trp* состоит из пяти протяженных структурных генов, которому предшествует регион, включающий в себя регуляторные участки (промотор, оператор, участок лидерного пептида, аттенуатор)

РНК-полимераза транскрибирует кодирующий регион. *Данный механизм антитерминации зависит от внешних факторов, способных влиять на трансляцию рибосомой лидирующей последовательности.*

Пять структурных генов оперона *trp* расположены рядом друг с другом и кодируют три фермента, превращающих хоризмовую кислоту в триптофан. Как показано на , транскрипция начинается с промотора,

расположенного с левого конца кластера. Экспрессия *trp*-оперона регулируется при помощи двух независимых механизмов. Репрессия триптофановых генов осуществляется репрессорным белком (кодируемым геном *trpR*), который связывается с прилегающим к промотору оператором. Аттенуация контролирует работу РНК-полимеразы, определяя, произойдет ли терминация в участке, предшествующем первому структурному гену.



Аттенуатор контролирует движение РНК-полимеразы вдоль последовательности генов оперона *trp*. Транскрипция инициируется на промоторе и продолжается до положения 90, где фермент, перед тем как пройти аттенуатор, делает паузу в положении 140. В отсутствие триптофана полимераза транскрибирует структурные гены (*trpE* начинается в положении +163). При отсутствии триптофана вероятность терминации с высвобождением 140-нуклеотидной лидерной последовательности РНК достигает ~90%

Впервые явление аттенуации было отмечено в экспериментах, показавших, что удаление последовательности, расположенной между оператором и кодирующей областью *trpE*, приводит к повышенной экспрессии структурных генов. Этот эффект не зависел от репрессии: увеличивались как фоновый, так и дерепрессированный уровень экспрессии. Следовательно, этот сайт оказывал влияние на события, происходящие уже *после* того, как РНК-полимераза сдвигалась с промотора и начинала транскрипцию (независимо от условий, в которых происходила инициация).

Аттенуатор (ρ-независимый терминатор) расположен между промотором и геном *trpE*. Он выступает в роли барьера, препятствующего транскрипции структурных генов. РНК-полимераза терминирует транскрипцию в области аттенуатора *in vivo* и *in vitro*, образуя при этом транскрипт длиной 140 пн.

Как показано на рис. 13.7, терминация транскрипции на аттенуаторе зависит от уровня триптофана в клетке. В присутствии достаточных количеств триптофана, терминация транскрипции триптофановых генов происходит эффективно. В отсутствие триптофана РНК-полимераза преодолевает аттенуатор и транскрибирует структурные гены.

Репрессия и аттенуация тоже зависят от содержания триптофана в клетке. В присутствии триптофана оперон репрессирован, и большая часть молекул РНК-полимеразы, уже переместившихся с промотора, терминируется в аттенуаторе. При удалении триптофана РНК-полимераза получает свободный доступ к промотору, и процесс транскрипции не прерывается преждевременно.

Аттенуация может изменять эффективность транскрипции примерно в 10 раз. При эффективной терминации около 10% молекул РНК-полимеразы, достигающих аттенуатора, способны преодолеть его. В условиях аттенуации почти все молекулы РНК-полимеразы способны продолжать транскрипцию. Наряду с 70-кратным увеличением эффективности инициации в результате снятия репрессии аттенуация делает возможным 700-кратное изменение уровня активности оперона.

13.5 Аттенуация контролируется с помощью трансляции

Ключевые понятия:

- Лидерная последовательность *trp*-оперона представлена открытой рамкой считывания, состоящей из 14 кодонов, два из которых являются кодонами триптофана.
- Структура РНК в аттенуаторе зависит от того, транслируется ли эта рамка считывания.
- В присутствии триптофана лидерная последовательность транслируется в белок, в области аттенуатора формируется шпилька, и транскрипция терминируется.
- В отсутствие триптофана рибосома останавливает трансляцию, достигнув триптофановых кодонов, а формирующаяся при этом альтернативная вторичная структура мРНК препятствует образованию шпильки и транскрипция продолжается.

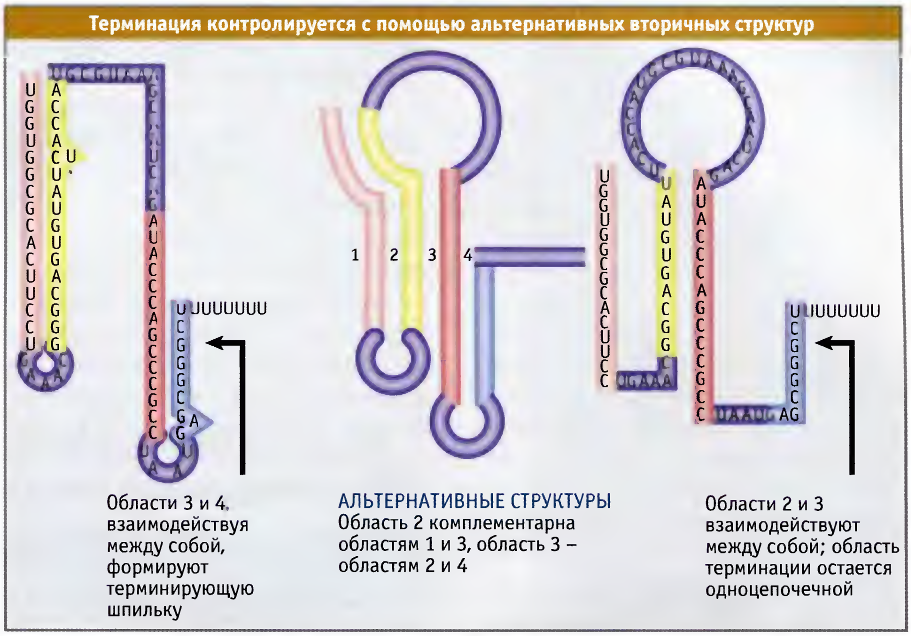


РИС. 13.8 Лидерная область *trp*-оперона может существовать в различных конформациях, образующихся при спаривании определенных оснований. В центре схематично изображены четыре области, которые могут взаимодействовать между собой. Область 1 комплементарна области 2, область 2 — области 3, а область 3, в свою очередь, — области 4. Слева представлена конформация, образующаяся при спаривании области 1 с областью 2, а области 3 — с областью 4. Справа — конформация, образующаяся при взаимодействии областей 2 и 3, при этом области 1 и 4 остаются неспаренными

Каким образом содержание триптофана в клетке влияет на терминацию транскрипции в аттенуаторе? Данные о нуклеотидной последовательности лидерной области позволяют предположить механизм этого процесса. Короткая кодирующая последовательность этого региона транслируется в **лидерный пептид**, состоящий из 14 аминокислот. Как показано на рис. 13.6, в эту область входит сайт связывания с рибосомой, в котором за AUG-кодоном следует участок из 13 кодонов, содержащий два остатка триптофана. При истощении запасов триптофана в клетке рибосомы инициируют трансляцию лидерного пептида, но останавливают ее, когда достигают триптофановых кодонов. Исходя из нуклеотидной последовательности мРНК, можно предположить, что такая **задержка рибосом**, в свою очередь, может влиять на терминацию в аттенуаторе.

Лидерная последовательность может быть представлена в виде структур, образующихся при альтернативных вариантах спаривания оснований. Способность рибосомы проходить через лидерную область может регулироваться с помощью перехода из одной структуры в другую. Именно от этой структуры зависит, обеспечит ли мРНК условия, необходимые для терминации.

На РИС. 13.8 изображены эти структуры. У первой область 1 спарена с областью 2, и область 3 — с областью 4. Спаривание областей 2 и 4 приводит к формированию шпильки, предшествующей последовательности U_8, что является необходимым сигналом для терминации. Вероятно, РНК приобретает такую структуру без какого-либо внешнего влияния.

В том случае, если спаривание областей 1 и 2 невозможно, образуется другая структура. Она возникает при взаимодействии области 2 с областью 3. Область 4 при этом теряет партнера для спаривания и остается одноцепочечной. В результате образование терминирующей шпильки невозможно.

Как показано на РИС. 13.9, положение рибосомы может определять, какая из структур будет образована, причем таким образом, чтобы терминация аттенуировалась только в отсутствие триптофана. Критическим параметром в этом случае будет положение триптофановых кодонов в кодирующей последовательности лидерного пептида.

В присутствии триптофана рибосомы способны синтезировать лидерный пептид. Они продолжают передвигаться вдоль лидерной области мРНК к кодону UGA, расположенному между регионами 1 и 2. Как показано в нижней части рисунка, продвигаясь к этой точке, рибосомы захватывают регион 2 и предотвращают его спаривание. В результате, регион 3 получает возможность спариться с регионом 4 и образовать терминирующую шпильку. При таких условиях РНК полимераза терминирует транскрипцию в аттенуаторе.

Как показано в верхней части рисунка, в отсутствие триптофана рибосома задерживается на trp-кодонах, входящих в состав области 1. При этом область 1 изо-

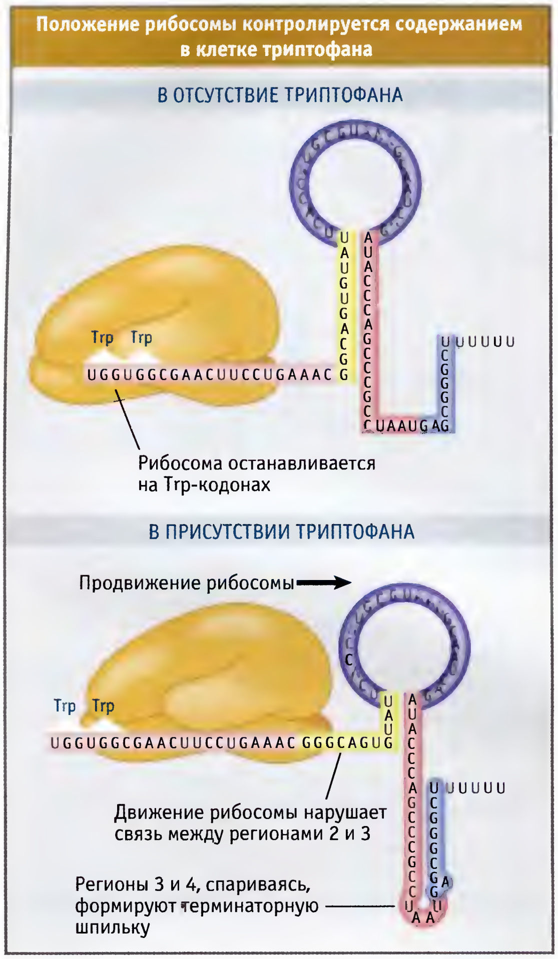


Рис. 13.9 То, какая именно вторичная структура образуется в аттенуаторе, зависит от положения рибосомы, которая определяет, провзаимодействуют ли регионы 3 и 4 с образованием терминаторной шпильки

лируется рибосомой и не может провзаимодействовать с областью 2. Это означает, что области 2 и 3 должны спариться раньше, чем область 4 начнет транскрибироваться. В результате область 4 остается в одноцепочечной форме. В отсутствие терминирующей шпильки РНК-полимераза продолжает транскрипцию, минуя аттенуатор.

Четкая координация последовательности регуляторных событий — необходимое условие аттенуации. Для того чтобы рибосома могла контролировать формирование необходимой для терминации вторичной структуры мРНК, *РНК-полимераза должна достигать сайта терминации в момент трансляции лидерной последовательности.* Для правильного согласования всех событий необходим сайт, вызывающий остановку РНК-полимеразы в положении, соответствующем 90-му основанию лидерной последовательности. РНК-полимераза не возобновляет своей работы до тех пор, пока рибосома не закончит трансляцию лидерного пептида. После этого фермент высвобождается и сдвигается к сайту аттенуации. К тому времени, как полимераза окажется там, вторичная структура области аттенуации уже будет сформирована.

На показано, какую роль в регуляции экспрессии триптофанового оперона играет Trp-тРНК. *Обеспечивая механизм, реагирующий на неадекватность снабжения Trp-тРНК, аттенуация отвечает непосредственно на потребность клетки в триптофане.*

Насколько широко используется явление аттенуации в качестве регуляторного механизма экспрессии бактериальных оперонов? Оно используется по крайней мере в шести оперонах, кодирующих ферменты, задействованные в биосинтезе аминокислот. Следовательно, механизм обратной связи между содержанием аминокислот, необходимых для белкового синтеза (что отражается в количестве доступных аминоацил-тРНК), и образованием ферментов, возможно, имеет общий характер.

При включении рибосомы в контроль вторичной структуры РНК в ответ на уровень доступной аминоацил-тРНК устанавливается обратная связь между присутствием аминоацил-тРНК и уровнем транскрипции оперона. Можно сказать, что аминоацил-тРНК выступает в роли корепрессора транскрипции. Регуляторный механизм опосредуется изменениями в конформации дуплексных регионов мРНК; явление аттенуации является ярким сви-

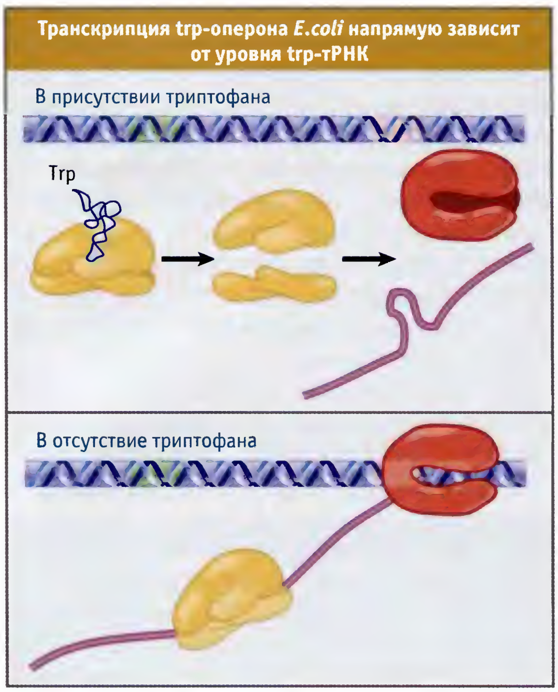


Рис. 13.10 В присутствии trp-тРНК рибосома транслирует лидирную последовательность и высвобождается. При этом формируется терминирующая шпилька, в области которой РНК-полимераза терминирует транскрипцию. В отсутствие trp-тРНК рибосома блокируется, терминирующая шпилька не образуется и РНК-полимераза продолжает транскрипцию

детельством важности вторичной структуры молекулы РНК для осуществления терминации.

В клетках *E. coli* и *B. subtilis* используются одни и те же типы регуляторных механизмов, в которых структура мРНК зависит от присутствия или отсутствия тРНК. Точные механизмы индивидуальных взаимодействий, однако, могут различаться. Конечный результат одинаков: ингибирование продукции ферментов синтеза при избытке триптофана и активация при его дефиците, который проявляется в накоплении свободной тРНК^{Trp}.

13.6 Инактивация экспрессии генов с помощью антисмысловой РНК

Ключевые понятия:

- Антисмысловые гены, при введении их в эукариотические клетки, блокируют экспрессию генов-мишеней

Некоторые РНК могут контролировать активность других РНК с помощью механизма комплементарного спаривания оснований. Как для про-, так и для эукариот описано множество примеров, в которых одноцепочечная РНК (обычно очень короткая), связываясь с комплементарным участком мРНК, блокирует ее экспрессию. Одними из самых первых иллюстраций этого эффекта стали эксперименты, в которых в эукариотические клетки искусственно вводили **антисмысловые гены**.

Антисмысловые гены были сконструированы путем изменения ориентации гена по отношению к его промотору. Как изображено на рис. 13.11, в результате транскрибировалась антисмысловая цепь РНК-матрицы. Антисмысловая РНК, связываясь с РНК-мишенью, инактивировала ее. Это явление характерно как для про-, так и для эукариотических клеток. Антисмысловая РНК служит искусственным регулятором активности транскриптов-мишеней. Антисмысловой ген тимидинкиназы ингибирует экспрессию исходного гена тимидинкиназы. Достоверная количественная оценка этому эффекту не дана, однако предполагается, что для его проявления необходим избыток антисмысловой РНК (возможно, существенный избыток).

На каком уровне антисмысловая РНК ингибирует экспрессию генов? Свой эффект антисмысловая РНК может проявлять на стадии транскрипции исходного гена, процессинга РНК-продукта, или трансляции транскриптов. Эксперименты с разными модельными системами показали, что ингибирование экспрессии гена-мишени зависит от процесса формирования РНК-РНК дуплекса, однако может происходить как в ядре, так и в цитоплазме. В случае когда антисмысловой ген стабильно присутствует в культуре клеток, РНК-дуплекс, состоящий из смысловой и антисмысловой цепей, формируется в ядре. В случае инъекции антисмысловой РНК в цитоплазму, трансляция ингибируется в результате формирования дуплекса между антисмысловой РНК и 5′-регионом мРНК-мишени.

Данный эффект лежит в основе метода «выключения» генов; к примеру, функция регуляторного гена может быть изучена с помощью введения в клетку его антисмыслового варианта. Расширить возможности данной техники можно, поместив антисмысловой ген под контроль промотора, который сам по себе является регулируемым элементом. Воздействуя на промотор и контролируя тем самым продукцию антисмысловой РНК, можно включать и выключать ген-мишень. Данный подход позволяет оценить значение продолжительности экспрессии гена-мишени.

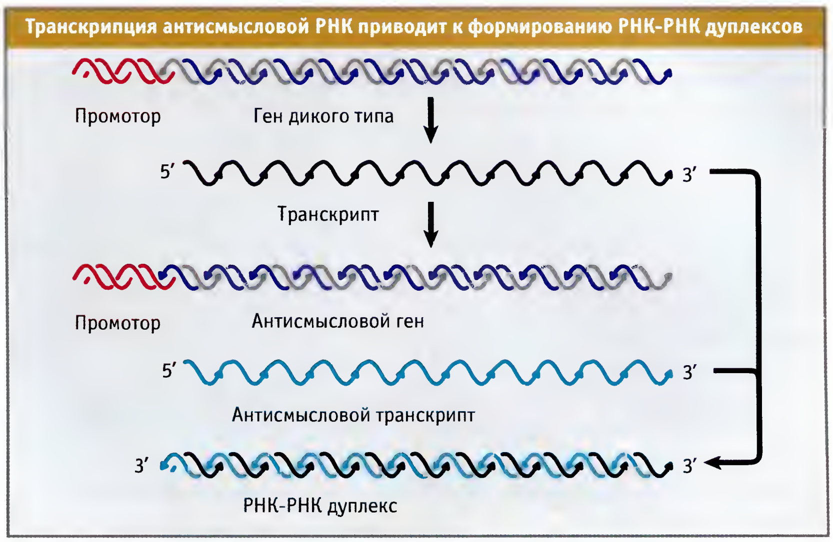


Рис. Антисмысловая РНК образуется в результате изменения ориентации гена по отношению к промотору и способна взаимодействовать с транскриптами дикого типа с образованием двухцепочечной структуры

13.7 Регуляция трансляции с помощью молекул малых РНК

Ключевые понятия:

- Регуляторная РНК осуществляет свои функции, формируя дуплекс с РНК-мишенью.
- Дуплекс может блокировать инициацию транскрипции, вызывать терминацию транскрипции, или служить мишенью для действия эндонуклеаз.

Репрессоры и активаторы являются *транс*-активными белками, однако в формировании регуляторной сети, помимо белков, могут участвовать и молекулы РНК. Первоначальная модель организации оперона не давала ответа на вопрос о природе регуляторных молекул. Для того чтобы имитировать действие природных регуляторных молекул РНК, которым в механизмах контроля придается все большее значение, были сконструированы синтетические антисмысловые РНК.

Как и белковый регулятор, малая регуляторная молекула РНК представляет собой независимо синтезируемую молекулу, которая диффундирует к специфической последовательности-мишени. Мишенью для действия регуляторной РНК является одноцепочечная нуклеотидная последовательность. Свой эффект регуляторная РНК реализует, связываясь с комплементарной ей последовательностью-мишенью и формируя двухцепочечную структуру.

Мы можем предложить два основных механизма действия регуляторной РНК:

- Образование дуплекса с последовательностью-мишенью либо блокирует ее способность выступать в роли специфического сайта связывания для других регуляторов, либо, напротив, способствует формированию нового сайта связывания. На РИС. 13.12 изображена ситуация, в которой белок, связывающийся с одноцепочечной РНК, не может осуществить свою функцию в случае формирования дуплекса между последовательностью-мишенью и регуляторной РНК. На РИС. 13.13 изображена противоположная ситуация, когда в результате образования двухцепочечной структуры формируется сайт, узнаваемый эндонуклеазой, разрушающей РНК-мишень.
- Формирование дуплекса в одной части молекулы-мишени изменяет конформацию других регионов, опосредованно влияя на их функцию. На РИС. 13.14 приведен пример такого влияния. Механизмы этого взаимодействия похожи на механизмы, которые обусловливают участие вторичной структуры РНК в аттенуации (см. разд. 13.2 «Аттенуация контролируется с помощью альтернативных вторичных структур»). Разница заключается лишь в том, что взаимодействующие регионы расположены на разных молекулах РНК, а не являются частями одной молекулы.

Общая черта двух типов РНК-опосредованной регуляции заключается в том, что активность последовательности-мишени зависит от ее вторичной структуры.

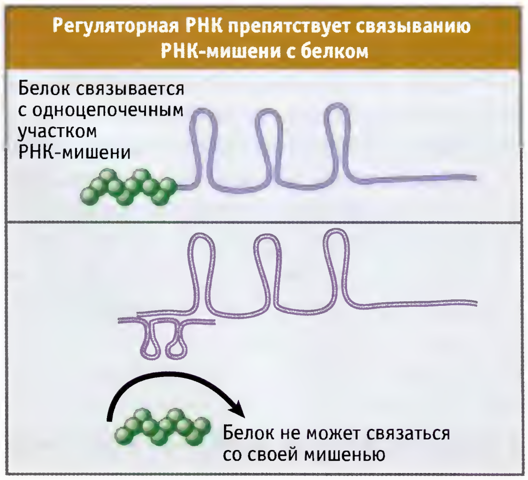


РИС. 13.12 Связывание регуляторной РНК с РНК-мишенью препятствует ее взаимодействию с белком

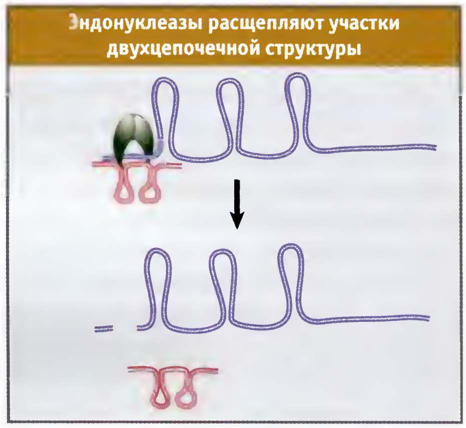


РИС. 13.13 В результате связывания регуляторной РНК и РНК-мишени формируется двухцепочечная структура, представляющая собой сайт для атаки нуклеазами

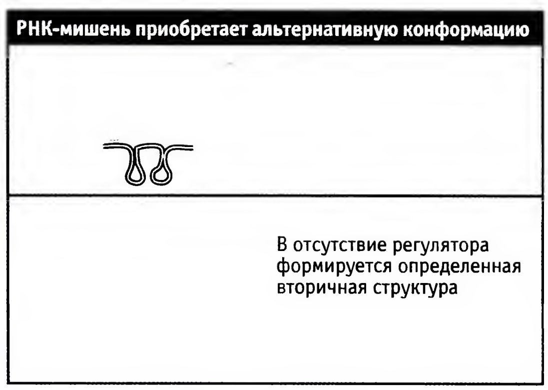


РИС. 13.14 Вторичная структура, образующаяся при спаривании оснований между двумя регионами РНК-мишени, не может сформироваться при взаимодействии этой РНК с регуляторной РНК. В описанном примере регуляторная РНК блокирует способность 3′-концевого участка молекулы РНК-мишени взаимодействовать с 5′-концевым участком

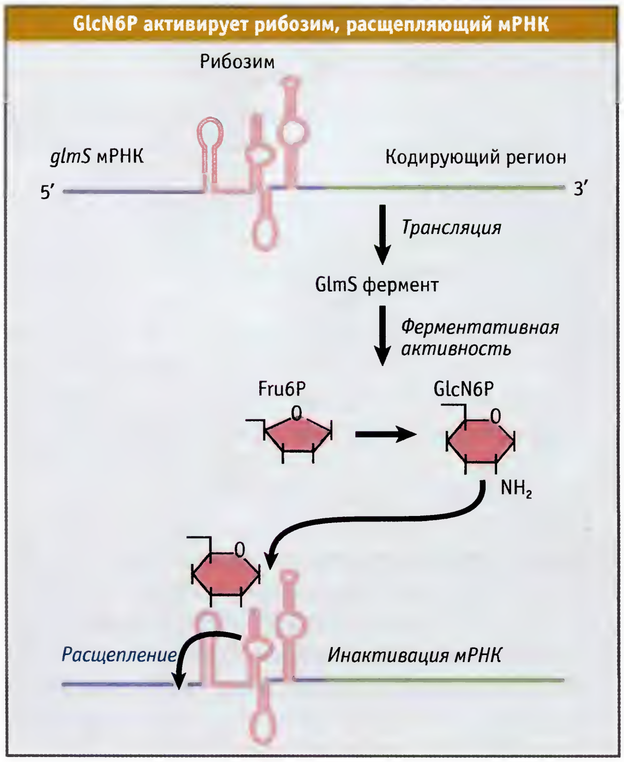


РИС. 13.15 5′-нетранслируемый регион мРНК фермента, ответственного за синтез GlcN6P, содержит в своем составе рибозим, активируемый продуктом метаболизма. Расщепляя свою собственную мРНК, рибозим инактивирует ее

Содержание малых регуляторных РНК может быть увеличено в результате стимуляции транскрипции кодирующего ее гена и снижено с помощью ферментов, осуществляющих деградацию РНК-продукта. Контролировать активность малых регуляторных РНК другими способами обычно не представляется возможным. Действительно, считалось, что РНК не обладает способностью к аллостерическим взаимодействиям; в отличие от репрессорных белков, контролирующих опероны, РНК не может отвечать на действие малых молекул изменением своей способности распознавать мишень.

Открытие **РНК-переключателей** (riboswitch) свидетельствует о том, что существуют исключения из этого правила. На РИС. 13.15 представлена система регуляции продукции метаболита — GlcN6P. Ген *glmS* кодирует фермент, синтезирующий GlcN6P (глюкозамин-6-фосфат) из фруктозо-6-фосфата и глутамина. Перед кодирующей областью в мРНК расположен протяженный 5′-нетранслируемый регион. Внутри UTR находится последовательность РНК, обладающая каталитической активностью, — рибозим (см. разд. 27.4 «Рибозимы с различной каталитической активностью»). В данном случае, обладая эндонуклеазной активностью, рибозим действует на свою собственную молекулу РНК. Он активируется при связывании с продуктом метаболизма — GlcN6P. Схема процесса следующая: активируясь при накоплении GlcN6P, рибозим расщепляет мРНК, делая невозможными последующие процессы трансляции. Прослеживается четкая параллель с аллостерическим контролем ак-

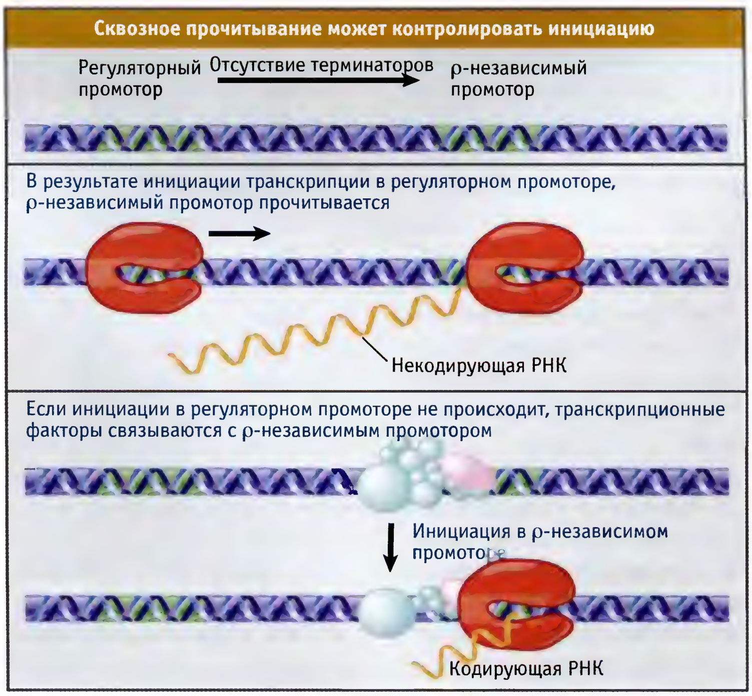


РИС. 13.16 Транскрипция, начавшаяся в промоторе, расположенном выше по ходу транскрипции, ингибирует инициацию транскрипции в промоторе, расположенном ниже. Прочитывая ρ-независимый промотор, РНК-полимераза не дает необходимым факторам терминации связаться с ним. РНК, транскрибируемая с вышележащего промотора, не выполняет кодирующей функции

тивности репрессорного белка конечными продуктами метаболического пути. Для бактерий описано несколько примеров таких РНК-переключателей.

Еще один регуляторный механизм заключается в транскрипции некодирующей последовательности РНК. Инициация в независимом промоторе супрессируется в результате транскрипции, начавшейся на другом промоторе, расположенном выше (против хода транскрипции). Пример такой ситуации изображен на РИС. 13.16. Инициируя транскрипцию на промоторе, расположенном выше, РНК-полимераза прочитывает независимый промотор, не связываясь с ним и не давая связаться с ним транскрипционным факторам. Этот эффект был описан в эукариотических системах и может являться следствием нарушения хромосомной структуры в области независимого промотора. РНК, транскрибируемая с другого (регуляторного) промотора, не выполняет кодирующей функции. Данное явление может объяснять присутствие в эукариотических ядрах некодирующих РНК-последовательностей, которые не выполняют регуляторных функций, но являются побочным продуктом функционирования систем контроля.

13.8 Регуляторные РНК бактерий

Ключевые понятия:

- **Бактериальные регуляторные РНК называют малыми РНК.**
- **Некоторые из бактериальных малых РНК находятся в связанном с белком Hfq состоянии, что повышает их эффективность.**
- **OxyS РНК на посттрансляционном уровне активирует или репрессирует экспрессию более чем 10 локусов.**

Регуляторные РНК бактерий представляют собой довольно короткие молекулы и в связи с этим получили общее название — малые РНК. В клетках *E. coli* содержится по крайней мере 17 разных малых РНК. Некоторые из этих малых РНК являются основными регуляторами клетки, воздействующими на множество генов. Свой эффект они осуществляют, связываясь с РНК-мишенями (как правило, мРНК), контролируя тем самым их стабильность или функцию.

Примером ситуации, в которой контролируется стабильность РНК-мишени, является взаимодействие антисмыслового регулятора RyhB с шестью мРНК, кодирующими белки, отвечающие за аккумуляцию железа в клетке *E.coli.* RyhB связывается с каждой из РНК-мишеней, формируя при этом двухцепочечную структуру, являющуюся субстратом для действия РНКазы Е. Особенность данного регуляторного мехнизма заключается в том, что регуляторная РНК при этом разрушается так же, как и мРНК.

Интересный механизм, в котором главную регуляторную роль играет РНК, включается при окислительном стрессе. На действие реактивных форм кислорода бактерия отвечает включением генов антиоксидантной системы. Пероксид водорода активирует белок OxyR, являющийся активатором транскрипции и контролирующий экспрессию нескольких индуцибельных генов. Один из этих генов — *oxyS*, кодирует малую РНК.

На РИС. 13.17 изображены две основные особенности механизмов контроля экспрессии *oxyS*. В нормальных условиях у бактерии дикого типа этот ген не экспрессируется. На гелях, изображенных в левой части рисунка, показано, что у мутантных бактерий с конститутивно активным геном *oxyR* ген *oxyS* экспрессируется на высоком уровне. Из этого следует, что ген *oxyS* активируется продуктом гена *oxyR*. На гелях справа показано, что OxyS РНК начинает транскрибироваться в течение минуты после добавления пероксида водорода.

OxyS РНК представляет собой короткую (109 нуклеотидов) некодирующую последовательность. Она является *транс*-активным регулятором, влияющим на экспрессию генов на посттрансляционных уровнях. OxyS РНК способна воздействовать более чем на 10 локусов-мишеней; экспрессию некоторых из них

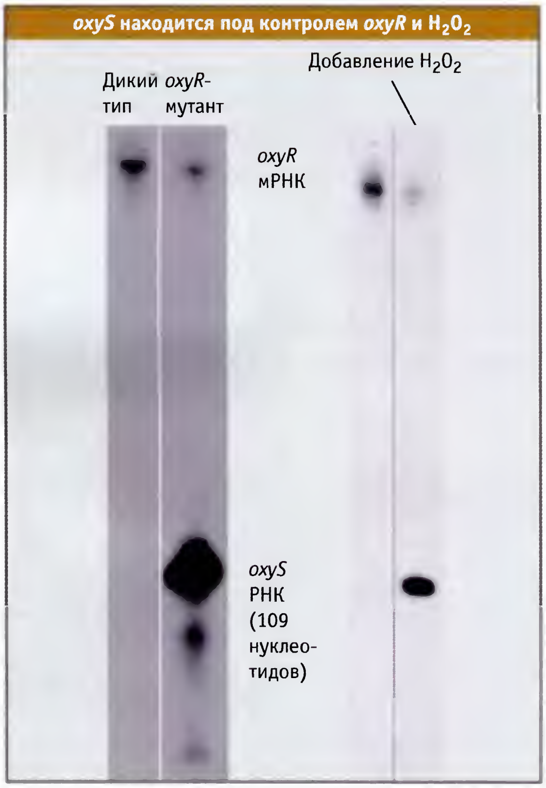


РИС. 13.17 На гелях слева показано, что транскрипция *oxyS* РНК индуцируется у *oxyR*-конститутивных мутантов. На гелях справа показано, что транскрипция *oxyS* РНК индуцируется в течение одной минуты после добавления пероксида водорода к культуре дикого типа. Воспроизведено из выпуска *Cell*, vol. 90, Altuvia, S., et al., A small stable RNA..., pp. 44–53. Защищено авторским правом в 1997 с разрешения Elsevier. Фотография любезно предоставлена Gisela Storz, National Institutes of Health

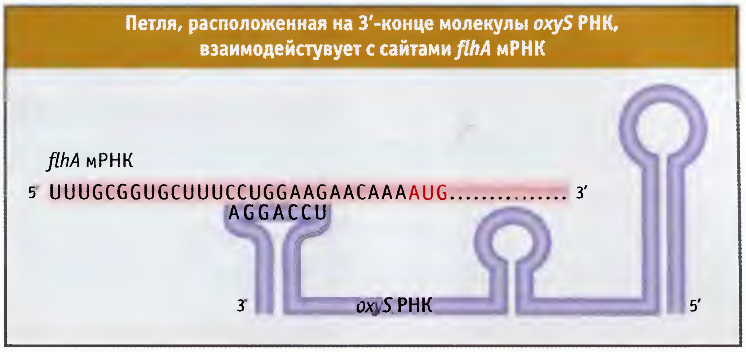


РИС. 13.18 *oxyS* РНК ингибирует трансляцию *flhA* мРНК, спариваясь с последовательностью, расположенной перед инициирующим кодоном AUG

она активирует, экспрессию других — репрессирует. На РИС. 13.18 показаны механизмы репрессии одной из мишеней — FlhA мРНК. Вторичная структура OxyS РНК содержит три структуры типа «стебель-петля». Петля, расположенная ближе к 3′-концу молекулы, комплементарна последовательности, предшествующей инициирующему кодону FlhA мРНК. Спаривание OxyS РНК с FlhA мРНК предотвращает связывание рибосомы с инициирующим кодоном и, таким образом, репрессирует трансляцию. Кроме того, свой вклад вносят и вторичные взаимодействия между OxyS РНК и кодирующей последовательностью FlhA.

Еще одной мишенью для действия *oxyS* является *rpoS* — ген, кодирующий альтернативный сигма-фактор (активирующийся в условиях стресса). Ингибируя продукцию сигма-фактора, *oxyS* обеспечивает условия, в которых специфический ответ на оксидативный стресс исключает запуск ответа, соответствующего другим стрессовым условиям. Ген *rpoS* регулируется еще двумя малыми РНК (DsrA и RprA), активирующими его. Три регуляторных РНК являются регуляторами с широким спектром возможностей и координируют ответы на самые разные внешние воздействия.

Действие всех трех малых РНК опосредовано РНК-связывающим белком — Hfq. Первоначально этот белок был идентифицирован как бактериальный хозяйский фактор, необходимый для репликации РНК бактериофага Qβ. Hfq напоминает эукариотические белки Sm, которые связываются с мяРНК (малые ядерные РНК), участвующими в регуляции экспрессии генов у эукариот (см. разд. 26.5 «мяРНК участвуют в сплайсинге»). Мутации в гене белка Hfq имеют множество проявлений, что позволяет определить этот белок как плейотропный. Hfq взаимодействует со многими малыми РНК *E. coli*. Усиливая способность OxyS РНК связываться с мРНК-мишенью, Hfq повышает эффективность ее действия. Свой эффект Hfq оказывает, вызывая небольшие изменения во вторичной структуре OxyS РНК, что способствует более прочному взаимодействию одноцепочечных участков OxyS РНК с мРНК-мишенью.

13.9 МикроРНК являются регуляторами во многих эукариотических организмах

Ключевые понятия:

- Геномы животных и растений кодируют большое количество коротких (~22 нуклеотида) молекул РНК, названных микроРНК.
- МикроРНК регулируют экспрессию генов, связываясь с комплементарными последовательностями мРНК-мишеней.

В эукариотических клетках регуляторами экспрессии генов служат очень короткие молекулы РНК. Впервые это явление было описано для нематод *Caenorhabditis elegans* при изучении взаимодействия между регуляторным геном *lin4* и его геном-мишенью *lin14*. На РИС. 13.19 изображена схема функционирования этой регуляторной системы. Ген *lin14* регулирует развитие личинки. Экспрессия гена *lin14* контролируется геном *lin4*, который транскрибируется с образованием коротких РНК длиной ~ 22 нуклеотида. Транскрипты *lin4* комплементарны 10-нуклеотидной последовательности, семикратно повторяющейся в 3′-нетранслируемом регионе *lin14*. Активность гена *lin4* ингибирует экспрессию *lin14* на посттранскрипционном уровне. Взаимодействие между регуляторной РНК и мРНК ведет к деградации мРНК. Интересной особенностью этого механизма является использование в качестве специфического сайта связывания с регулятором 3′-конца молекулы мРНК.

РНК *lin4* являет собой пример **микроРНК**. В геноме *C. elegans* присутствует примерно 80 генов, кодирующих микроРНК длиной 21–24 нуклеотида. Эти гены меняют профиль своей экспрессии по мере развития организма и, по всей видимости, являются регуляторами экспрессии других генов. Многие из микроРНК *C. elegans* входят в состав большой (15S) рибонуклеопротеиновой частицы.

Многие микроРНК *C. elegans* имеют гомологов у млекопитающих, а значит, такой механизм контроля экспрессии генов может быть распространен довольно широко. Кроме того, они найдены у растений. Из

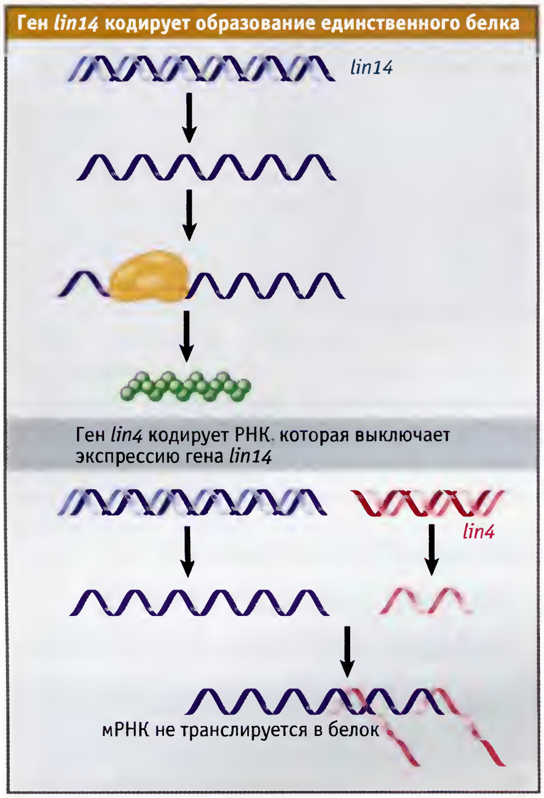


РИС. 13.19 *lin4* РНК регулирует экспрессию гена *lin14*, взаимодействуя с 3′-нетранслируемым участком *lin14* мРНК

шестнадцати микроРНК, обнаруженных у *Arabidopsis*, восемь (полностью идентичных) найдено у риса, что свидетельствует о консервативности данного регуляторного механизма.

Вирус SV40 кодирует образование микроРНК, комплементарных вирусным мРНК, продуцируемым на ранних стадиях вирусной инфекции. МикроРНК транскрибируются в вирусном цикле позже. Связываясь с ранними мРНК, микроРНК способствуют их деградации на той стадии развития вируса, когда необходимость в них уже отпала.

Механизм продукции микроРНК также консервативен и широко распространен. В случае с *lin4* ген транскрибируется с образованием транскриптов, которые формируют двухцепочечную структуру, являющуюся субстратом для действия нуклеазы Дайсер (Dicer). Фермент обладает N-концевой хеликазной активностью, за счет которой он расплетает РНК-дуплекс, и двумя нуклеазными доменами, обеспечивающими сходство с бактериальной рибонуклеазой III. Похожие ферменты обнаружены у мух, червей и растений. В результате расщепления первичного транскрипта образуется активная микроРНК. Нарушения в работе фермента ведут к блокировке продукции микроРНК и вызывают дефекты развития.

Другая стадия формирования микроРНК была охарактеризована у растений. В организме растений 3′-концевой нуклеотид микроРНК метилирован по рибозному остатку метилтрансферазой HEN1. Метилирование стабилизирует микроРНК.

13.10 РНК-интерференция — причина молчания генов

Ключевые понятия:

- В результате РНК-интерференции запускается процесс деградации мРНК, комплементарной любой из цепей короткой двухцепочечной регуляторной РНК (дцРНК).
- Двухцепочечная РНК способна подавлять экспрессию генов, приводя к их молчанию.

Регуляция активности мРНК с помощью микроРНК имеет одну природу с феноменом **РНК-интерференции**. Явление интерференции было открыто после того, как обнаружилось, что смысловая и антисмысловая РНК одинаково эффективны в ингибировании экспрессии генов. Причина этого заключается в том, что конструирование одноцепочечной РНК любого типа (предположительно) сопровождается контаминацией небольшими количествами двухцепочечной РНК (дцРНК).

Работа с системами *in vitro* показала, что дцРНК деградирует при АТФ-зависимом расщеплении, образуя при этом олигонуклеотиды длиной 21–23 основания, иногда называемые малыми интерферирующими РНК (миРНК). Как показано на РИС. 13.20, разрывы вносятся с 3′-конца длинной дцРНК, при этом образуются миРНК-фрагменты с короткими (два основания) выступающими 3′-концами. Тот же фермент (Дайсер), что отвечает за образование микроРНК, осуществляет и расщепление дцРНК на миРНК.

РНК-интерференция проявляется на посттранскрипционном уровне, когда миРНК индуцируют разрушение комплементарной мРНК. Как показано на РИС. 13.21, миРНК связывается с мРНК, комплементарной одной или обеим цепям минРНК, запуская тем самым процесс ее деградации с помощью нуклеаз. Вероятно, участие хеликаз в этом процессе необходимо. Расщепление мРНК происходит в центре спаренного с миРНК сегмента. Реакция протекает внутри рибонуклеопротеинового комплекса, названного RISC (RNA-induced silencing complex). Белки семейства Аргонавт (Argonaute) являются компонентами этого комплекса и обеспечивают расщепление мРНК. Метилирование 3′-концевой рибозы, возможно, является необходимым условием встраивания в комплекс миРНК.

До сих пор точно не выяснено, каким образом RISC-комплекс выключает экспрессию гена (см. разд. 31.1 «Гетерохроматин зависит от взаимодействия ДНК с гистонами»). Для реализации РНК-интерференции требуется участие РНК-полимеразы. Однако непонятно, влияет фермент на процесс РНК-интерференции напрямую (поскольку участвует в образовании необходимых транскриптов) или опосредованно (поскольку фермент участвует в связывании интерферирующего комплекса с транскриптами, а также в расплетании цепей ДНК во время транскрипции, способствуя тем самым взаимодействию с интерферирующим комплексом).

РНК-интерференция является эффективным подходом к выключению экспрессии специфических ге-

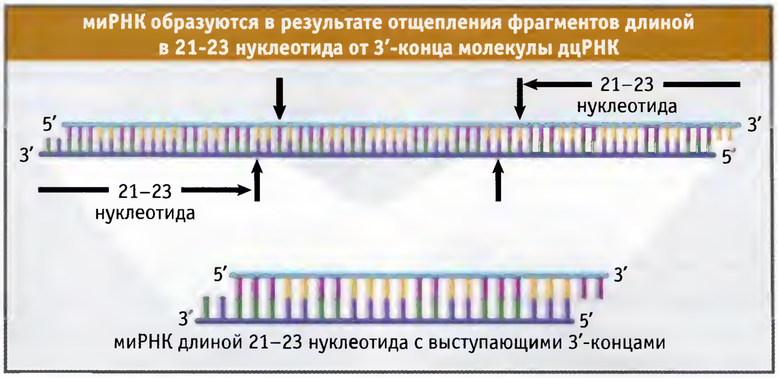


РИС. 13.20 миРНК, опосредующая явление РНК-интерференции, образуется в результате расщепления дцРНК на более короткие фрагменты. В ходе реакции с 3′-конца молекулы отщепляются фрагменты длиной 21–23 нуклеотида. Образующиеся в результате миРНК имеют выступающие 3′-концы

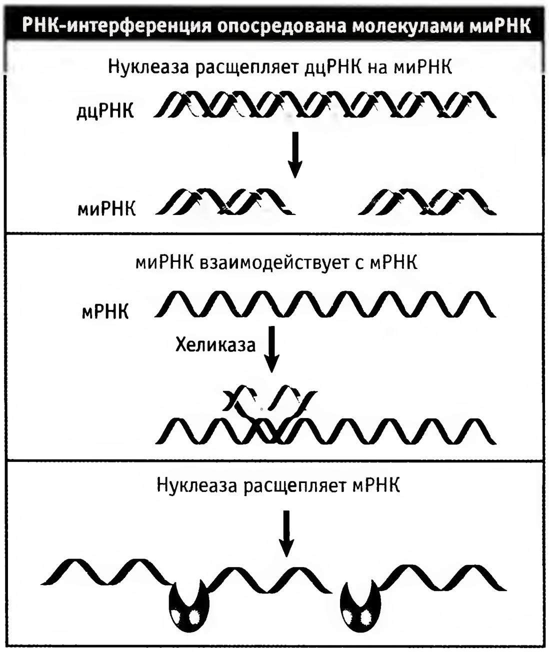


РИС. 13.21 Короткие фрагменты, образующиеся в результате расщепления более длинных молекул дцРНК, осуществляют деградацию соответствующей мРНК, обеспечивая тем самым явление РНК-интерференции

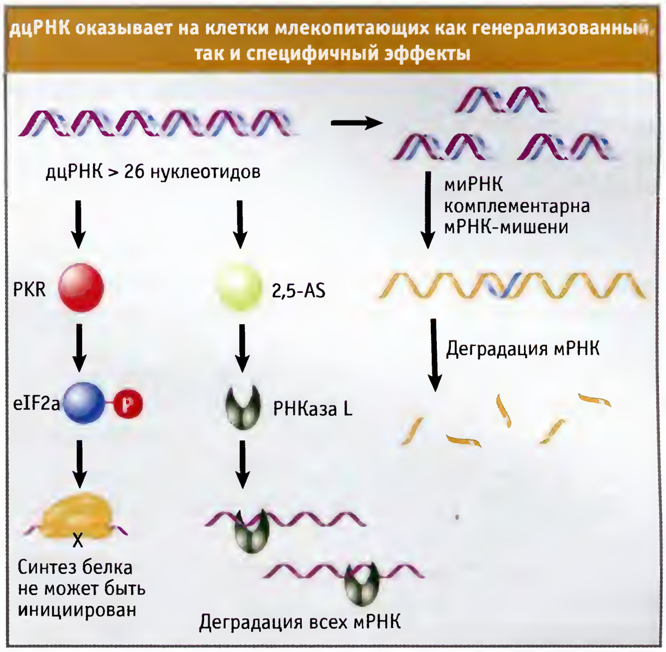


РИС. 13.22 дцРНК ингибируют синтез белков и запускают процесс деградации не только отдельных мРНК, но и всех остальных мРНК в клетках млекопитающих

нов-мишеней в клетках беспозвоночных (*C. elegans* и *Drosophila melanogaster*). Однако применение этого метода в клетках млекопитающих ограничено, где эффект дцРНК носит более генерализованный характер, приводя к деградации мРНК и остановке синтеза белка. На РИС. 13.22 изображено, какие процессы запускаются в клетках млекопитающих при действии дцРНК. дцРНК активирует фермент PKR, который инактивирует фактор инициации трансляции eIF2a (фосфорилируя его). Кроме того, он активирует 2,5-олигоаденилатсинтетазу, что приводит к активации РНКазы L, осуществляющей деградацию всех мРНК клетки. Эти реакции требуют участия дцРНК, длина которых превышает 26 нуклеотидов. При введении в клетки млекопитающих более коротких мРНК (21–23 нуклеотида) запускается деградация лишь отдельных мРНК, комплементарных дцРНК (точно так же, как и при использовании явления РНК-интерференции в клетках червей или мух). Учитывая этот факт, можно предположить, что вскоре РНК-интерференция станет универсальным механизмом для выключения экспрессии отдельных генов.

Примером экспериментов, в которых успешно используется явление РНК-интерференции, может служить систематический анализ экспрессии генов у *C. elegans*.

Эффект потери функции достигается при кормлении червей бактериями, экспрессирующими дцРНК, гомологичные гену-мишени. С помощью библиотеки бактерий, в которой бактерии экспрессируют дцРНК, соответствующие разным генам, черви были проанализированы на предмет последствий от выключения большей части (86%) генов.

РНК-интерференция имеет одну природу с естественным процессом выключения генов. Клетки растений и грибов демонстрируют **молчание РНК** (иногда называемое посттранскрипционным молчанием генов), при котором экспрессия гена ингибируется дцРНК. Самым распространенным источником дцРНК служит реплицирующийся вирус. Данный механизм может выступать в качестве защитного барьера против вирусной инфекции. дцРНК, образующиеся при инфицировании вирусом растительной клетки, запускают процесс ингибирования экспрессии растительного генома. Характерной особенностью феномена молчания РНК является то, что наблюдается он не только в инфицированных вирусом клетках, но распространяется в организме растения системно. Предполагают, что передача сигнала представляет собой процесс перемещения РНК или фрагментов РНК. Возможно, в этом участвуют те же механизмы, которые обеспечивают движение самого вируса. Вероятно, для реализации молчания РНК необходимо увеличение сигнала, происходящее в результате РНК-зависимого синтеза РНК. При этом полимераза использует миРНК в качестве затравок для синтеза на матрице комплементарной им РНК.

Похожим процессом является **косупрессия**. При косупрессии введение в клетку трансгена приводит к выключению экспрессии исходного гена. Трансген в этом случае кодирует образование как антисмысловых, так и смысловых РНК, которые ингибируют экспрессию собственного гена клетки.

Молчание генов реализуется посредством РНК-РНК взаимодействий. Возможно, дцРНК способны ингибировать экспрессию генов, взаимодействуя с ДНК. Если в растительный геном ввести ДНК, гомологичную последовательности вирусной РНК, то при репликации РНК вируса происходит метилирование этой ДНК. Из этого следует, что РНК-последовательность индуцирует метилирование ДНК-последовательности. Похожим образом происходит метилирование растительной ДНК, комплементарной дцРНК. Способствуя репрессии транскрипции, метилирование ДНК может служить другим инструментом достижения молчания отдельных генов (см. разд. 24.18 «Экспрессия генов связана с деметилированием»).

13.11 Резюме

Регуляция экспрессии генов может быть позитивной (и осуществляться с помощью факторов, активирующих ген) и негативной (происходящей при участии факторов, репрессирующих ген). Первым и наиболее часто используемым уровнем контроля является инициация транскрипции. Стадия терминации транскрипции также подвергается контролю. Трансляция контролируется регуляторами, взаимодействующими с мРНК. В качестве регуляторов могут выступать белки или РНК. Свою регуляторную функцию белки часто осуществляют с помощью аллостерических взаимодействий, а регуляторная РНК изменяет вторичную структуру РНК-мишени, спариваясь с ней по принципу комплементраности. Организация регуляторных сетей подразумевает связь регуляторов между собой: продукция и активность одних контролируется другими.

Аттенуация — механизм регуляции транскрипции бактериальных оперонов, в основе которого лежит контроль терминации. Это явление часто используется для контроля активности оперонов, кодирующих ферменты биосинтеза аминокислот. Полицистронная мРНК таких оперонов начинается с последовательности, способной формировать альтернативные вторичные структуры. Одна из таких структур представляет собой шпильку, выступающую в роли ρ-независимого терминатора, располагающегося перед структурными генами; альтернативные структуры шпильки не формируют. Разные типы взаимодействий определяют, будет ли образована шпилька. Примером таких взаимодействий может служить взаимодействие белка с мРНК, препятствующее формированию альтернативных вторичных структур. Для *trp*-оперона *B. subtilis* эту функцию выполняет белок TRAP; он контролируется белком анти-TRAP, чья продукция, в свою очередь, контролируется уровнем свободной аминоацил-тРНК^{Trp}. То, какая именно вторичная структура будет сформировна в *trp*- и других оперонах *E. coli*, зависит от того, продолжится ли трансляция после короткой лидерной последовательности, включающей в себя кодоны аминокислоты (аминокислот), являющейся продуктом данной системы. В присутствии аминоацил-тРНК, нагруженной этой аминокислотой, рибосомы транслируют лидерный пептид, в результате чего формируется вторичная структура, обеспечивающая последующую терминацию. В отсутствие аминоацил-тРНК рибосома приостанавливает свою работу, и РНК упаковывается в альтернативную вторичную структуру, в которой необходимая для терминации шпилька отсутствует. Таким образом, уровень аминоацил-тРНК в клетке обратно пропорционален уровню биосинтеза соответствующей аминокислоты.

Малые регуляторные РНК обнаружены как у бактерий, так и у эукариот. В клетках *E. coli* присутствует порядка 17 различных малых РНК. *oxyS* РНК на посттранскрипционном уровне способна воздействовать более чем на 10 локусов-мишеней; экспрессию некоторых из них она активирует, экспрессию других — подавляет. Связываясь с мРНК-мишенью, малая РНК формирует двухцепочечную структуру, экранируя при этом сайт связывания с рибосомой, то есть репрессирует трансляцию. МикроРНК имеют длину ~22 нуклеотида и про-

дуцируются во многих эукариотических клетках путем расщепления более длинных транскриптов. Свою функцию они реализуют, спариваясь с мРНК-мишенью и формируя двухцепочечную структуру, которая деградирует под действием эндонуклеаз. В результате деградации мРНК экспрессия гена снижается. РНК-интерференция лежит в основе метода избирательной инактивации эукариотических генов. Метод заключается в индуцировании короткими дцРНК фрагментами, комплементарными мРНК-мишени, деградации мРНК. Возможно, этот феномен имеет одну природу с естественной защитной системой растений, названной молчанием РНК.

Литература

13.3 Терминация *trp*-генов у *Bacillus subtilis* контролируется уровнями триптофана и тРНК^{Trp}

Обзор

Gollnick, P. (1994). Regulation of the *B. subtilis trp* operon by an RNA-binding protein. *Mol. Microbiol.* 11, 991–997.

Исследование

Antson, A. A. et al. (1999). Structure of the *trp* RNA-binding attenuation protein, TRAP, bound to RNA. *Nature* 401, 235–242.

Babitzke, P. and Yanoksy, C. (1993). Reconstitution of *B. subtilis trp* attenuation *in vitro* with TRAP, the *trp* RNA-binding attenuation protein. *Proc. Natl. Acad. Sci. USA* 90, 133–137.

Otridge, J. and Gollnick, P. (1993). MtrB from *B. subtilis* binds specifically to *trp* leader RNA in a tryptophan-dependent manner. *Proc. Natl. Acad. Sci. USA* 90, 128–132.

Valbuzzi, A. and Yanofsky, C. (2001). Inhibition of the *B. subtilis* regulatory protein TRAP by the TRAP-inhibitory protein, AT. *Science* 293, 2057–2059.

13.4 Триптофановый оперон *E. coli* контролируется с помощью аттенуации

Обзор

Yanofsky, C. (1981). Attenuation in the control of expression of bacterial operons. *Nature* 289, 751–758.

13.5 Аттенуация контролируется с помощью трансляции

Обзор

Bauer, C. E., Carey, J., Kasper, L. M., Lynn, S. P., Waechter, D. A., and Gardner, J. F. (1983). Attenuation in bacterial operons. In Beckwith, J., Davies, J., and Gallant, J. A., eds. *Gene Function in Prokaryotes.* Cold Springs Harbor, NY: Cold Spring Harbor Press. pp. 65–89.

Landick, R. and Yanofsky, C. (1987). In Neidhardt, F. C., ed., E. coli *and* S. typhimurium: *Cellular and Molecular Biology.* Washington, D.C.: American Society for Microbiology. pp. 1276–1301.

Yanofsky, C. and Crawford, I. P. (1987). In Ingraham, J. L., et al., eds., Escherichia coli *and* Salmonella typhimurium, Washington, D.C.: American Society for Microbiology. pp. 1453–1472.

Исследование

Lee, F. and Yanofsky, C. (1977). Transcription termination at the *trp* operon attenuators of *E. coli* and *S. typhimurium*: RNA secondary structure and regulation of termination. *Proc. Natl. Acad. Sci. USA* 74, 4365–4368.

Zurawski, G. et al. (1978). Translational control of transcription termination at the attenuator of the *E. coli tryptophan* operon. *Proc. Natl. Acad. Sci. USA* 75, 5988–5991.

13.6 Инактивация экспрессии генов с помощью антисмысловой РНК

Исследование

Izant, J. G. and Weintraub, H. (1984). Inhibition of thymidine kinase gene expression by antisense RNA: a molecular approach to genetic analysis. *Cell* 36, 1007–1015.

13.7 Регуляция трансляции с помощью молекул малых РНК

Исследование

Johnson, A., and O'Donnell, M. (2005). Cellular DNA replicases: components and dynamics at the replication fork. *Annu. Rev. Biochem.* 74, 283–315.

Martens, J. A., Laprade, L., and Winston, F. (2004). Intergenic transcription is required to repress the *Saccharomyces cerevisiae* SER3 gene. *Nature* 429, 571–574.

Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A., and Breaker, R. R. (2004). Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* 428, 281–286.

13.8 Регуляторные РНК бактерий

Обзор

Gottesman, S. (2002). Stealth regulation: biological circuits with small RNA switches. *Genes Dev.* 16, 2829–2842.

Исследование

Altuvia, S., Weinstein-Fischer, D., Zhang, A., Postow, L., and Storz, G. (1997). A small, stable RNA induced by oxidative stress: role as a pleiotropic regulator and antimutator. *Cell* 90, 43–53.

Altuvia, S., Zhang, A., Argaman, L., Tiwari, A., and Storz, G. (1998). The *E. coli* OxyS regulatory RNA represses *fhlA* translation by blocking ribosome binding. *EMBO J.* 17, 6069–6075.

Massé, E., Escorcia, F. E., and Gottesman, S. (2003). Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli. Genes Dev.* 17, 2374–2383.

Moller, T., Franch, T., Hojrup, P., Keene, D. R., Bachinger, H. P., Brennan, R. G., and Valentin-Hansen, P. (2002). Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. *Mol. Cell* 9, 23–30.

Wassarman, K. M., Repoila, F., Rosenow, C., Storz, G., and Gottesman, S. (2001). Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev.* 15, 1637–1651.

Zhang, A., Wassarman, K. M., Ortega, J., Steven, A. C., and Storz, G. (2002). The Sm-like Hfq protein increases OxyS RNA interaction with target mRNAs. *Mol. Cell* 9, 11–22.

13.9 МикроРНК являются регуляторами во многих эукариотических организмах

Исследование

Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409, 363–366.

Ketting, R. F., Fischer, S. E., Bernstein, E., Sijen, T., Hannon, G. J., and Plasterk, R. H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans. Genes Dev.* 15, 2654–2659.

Lau, N. C., Lim, l.e. E. P., Weinstein, E. G., and Bartel, d.a. V. P. (2001). An abundant class of tiny RNAs with probable regulatory roles in *C. elegans. Science* 294, 858–862.

Lee, R. C. and Ambros, V. (2001). An extensive class of small RNAs in *C. elegans. Science* 294, 862–864.

Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14. Cell* 75, 843–854.

Mourelatos, Z., Dostie, J., Paushkin, S., Sharma, A., Charroux, B., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002). miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev.* 16, 720–728.

Reinhart, B. J., Weinstein, E. G., Pades, M. W., Bartel, B., and Bartel, D. P. (2002). MicroRNAs in plants. *Genes Dev.* 16, 1616–1626.

Sullivan, C. S., Grundhoff, A. T., Tevethia, S., Pipas, J. M., and Ganem, D. (2005). SV40-encoded microRNAs regulate viral gene expression and reduce susceptibility to cytotoxic T cells. *Nature* 435, 682–686.

Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans. Cell* 75, 855–862.

Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R. W., Steward, R., and Chen, X. (2005). Methylation as a crucial step in plant microRNA biogenesis. *Science* 307, 932–935.

Zamore, P. D., and Haley, B. (2005). Ribo-gnome: the big world of small RNAs. *Science* 309, 1519–1524.

13.10 РНК-интерференция – причина молчания генов

Обзор

Ahlquist, P. (2002). RNA-dependent RNA polymerases, viruses, and RNA silencing. *Science* 296, 1270–1273.

Matzke, M., Matzke, A. J., and Kooter, J. M. (2001). RNA: guiding gene silencing. *Science* 293, 1080–1083.

Schwartz, D. S. and Zamore, P. D. (2002). Why do miRNAs live in the miRNP? *Genes Dev.* 16, 1025–1031.

Sharp, P. A. (2001). RNA interference—2001. *Genes Dev.* 15, 485–490.

Tijsterman, M., Ketting, R. F., and Plasterk, R. H. (2002). The genetics of RNA silencing. *Annu. Rev. Genet.* 36, 489–519.

Исследование

Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 411, 494–498.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* 391, 806–811.

Hamilton, A. J. and Baulcombe, D. C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* 286, 950–952.

Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., Welchman, D. P., Zipperlen, P., and Ahringer, J. (2003). Systematic functional analysis of the *C. elegans* genome using RNAi. *Nature* 421, 231–237.

Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. (2004). Human argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol. Cell* 15, 185–197.

Mette, M. F., Aufsatz, W., van der Winden, J., Matzke, M. A., and Matzke, A. J. (2000). Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *EMBO J.* 19, 5194–5201.

Montgomery, M. K., Xu, S., and Fire, A. (1998). RNA as a target of double-stranded RNA-mediated genetic interference in *C. elegans. Proc. Natl. Acad. Sci. USA* 95, 15502–15507.

Ngo, H., Tschudi, C., Gull, K., and Ullu, E. (1998). Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei. Proc. Natl. Acad. Sci. USA* 95, 14687–14692.

Schramke, V., Sheedy, D. M., Denli, A. M., Bonila, C., Ekwall, K., Hannon, G. J., and Allshire, R. C. (2005). RNA-interference-directed chromatin modification coupled to RNA polymerase II transcription. *Nature* 435, 1275–1279.

Voinnet, O., Pinto, Y. M., and Baulcombe, D. C. (1999). Suppression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants. *Proc. Natl. Acad. Sci. USA* 96, 14147–14152.

Wassenegger, M., Heimes, S., Riedel, L., and Sanger, H. L. (1994). RNA-directed de novo methylation of genomic sequences in plants. *Cell* 76, 567–576.

Wateρuse, P. M., Graham, M. W., and Wang, M. B. (1998). Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proc. Natl. Acad. Sci. USA* 95, 13959–13964.

Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R. W., Steward, R., and Chen, X. (2005). Methylation as a crucial step in plant microRNA biogenesis. *Science* 307, 932–935.

Zamore, P. D., and Haley, B. (2005). Ribo-gnome: the big world of small RNAs. *Science* 309, 1519–1524.

Zamore, P. D., Tuschl, T., Sharp, P. A., and Bartel, D. P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* 101, 25–33.

Стратегии бактериофагов

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

- ДНК-связывающий сайт содержит два коротких α-спиральных региона, которые ложатся в соседние витки большой бороздки ДНК.
- ДНК-связывающий сайт — это несовершенная палиндромная последовательность из 17 пн.

14.12 «Спираль узнавания» задает специфичность к ДНК
- Взаимодействие репрессора с конкретными парами нуклеотидов в операторе определяется аминокислотной последовательностью «спирали узнавания».

14.13 Димеры репрессора связываются с оператором кооперативно
- Связывание репрессора с одним из операторов повышает вероятность посадки второго репрессора на соседний оператор.
- Сродство репрессора к O_L1 и O_R1 примерно на порядок выше, чем к другим операторам, что определяет их первоочередное связывание.
- Кооперативность позволяет репрессору связывать сайты *O*1–*O*2 при меньших концентрациях.

14.14 Находящийся на O_R2 репрессор взаимодействует с РНК-полимеразой на P_{RM}
- ДНК-связывающий регион репрессора на O_R2 взаимодействует с РНК-полимеразой и стабилизирует ее связывание с P_{RM}.
- Данный механизм является основой саморегуляции работы,репрессора.

14.15 Репрессор поддерживает цикл саморегуляции
- Связывание репрессора с O_L блокирует транскрипцию гена N с промотора P_L.
- Связывание репрессора с O_R блокирует транскрипцию гена *cro*, а также необходимо для транскрипции *cI*.
- Таким образом, связывание репрессора с операторами одновременно блокирует начало литического цикла и поддерживает его собственный синтез.

14.16 Кооперативный характер связывания повышает чувствительность регуляции
- Димеры репрессора, связавшие O_L1 и O_L2, взаимодействуют с димерами, связавшими O_R1 и O_R2; все четыре димера образуют октамер.
- Образование октамера приводит к сближению O_L3 с O_R3 и взаимодействию димеров репрессора.
- Кооперативный характер взаимодействия димеров репрессора с операторами повышает чувствительность регуляции.

14.17 Гены *cII* и *cIII* нужны для установления лизогении
- Продукты задержаннoранних генов *cII* и *cIII* необходимы, чтобы РНК-полимераза могла начать транскрипцию с промотора P_{RE}.
- Белок *cII* адаптирует промотор, а *cIII* защищает *cII* от деградации.
- Транскрипция с P_{RE} приводит к синтезу репрессора и мешает экспрессии *cro*.

14.18 Слабому промотору нужен белок *cII*
- Промотор P_{RE} имеет нестандартные последовательности на функционально-значимых участках –35 и –10.
- РНК-полимераза узнает этот промотор только в присутствии белка cII.
- Белок cII связывает участок –35 промотора P_{RE}.

14.19 Установлению лизогении предшествует ряд событий
- *cII* и *cIII* обеспечивают первичный всплеск синтеза репрессора, а также запускают ингибирование транскрипции поздних генов.
- С появлением репрессора транскрипция всех ранних генов выключается.
- Репрессор включает контур поддержания своего собственного синтеза.
- На заключительном этапе перехода в лизогенное состояние ДНК фага лямбда встраивается в геномную ДНК бактерии.

14.20 Без репрессора cro полный литический цикл невозможен
- Продукт гена *cro* связывается с теми же операторами, что и репрессор, но константы сродства у него другие.
- Если Cro связался с O_R3, он не дает РНК-полимеразе связаться с P_{RM} и, таким образом, блокирует самоподдержание репрессора.
- Если Cro связался с другими сайтами операторов O_R и O_L, он не дает РНК-полимеразе транскрибировать немедленноранние гены, что косвенным образом блокирует установление репрессора.

14.21 На чем основывается баланс между лизогенией и литическим циклом?
- Лизогения и литический цикл развиваются по одному сценарию — вплоть до момента одновременного присутствия Cro и лизогенного репрессора cI.
- Ключевой момент: удастся ли белку *cII* стимулировать наработку репрессора в количестве, достаточном для преодоления действия Cro?

14.22 Резюме

14.1 Введение

У некоторых фагов есть только одна стратегия выживания. Инфицировав бактерию, они подчиняют ее клеточные функции одной цели: продуцировать как можно большее число своих потомков, инфекционных частиц. Такой ход событий называют **литической инфекцией**, или литическим развитием; он приводит к гибели бактерии-хозяина. В типичном варианте литического развития геном фага (в форме ДНК или РНК) попадает внутрь бактерии, его гены транскрибируются в определенном порядке; генетический материал фага реплицируется; идет продукция белковых компонентов будущих инфекционных частиц. Цикл завершается процессом **лизиса**, который приводит к разрушению (лизированию) клетки бактерии-хозяина и высвобождению зрелых инфекционных частиц бактериофага (фаговых частиц).

Есть фаги, которые ведут двойное существование. В одних условиях они способны к продолжению рода через литический цикл, демонстрируя открытую стра-

тегию продукции как можно большего числа копий за минимальное время. Запасной же вариант стратегии бактериофага выражается в наличии у бактерии (в составе ее генома) латентной копии фага, так называемого **профага**. Эту форму существования и размножения бактериофагов называют **лизогенией**.

В лизогенных бактериях профаг встроен в геном и наследуется наравне с их собственными генами. Процесс превращения независимого генома бактериофага в профаг (линейную вставку в бактериальный геном) называют **интеграцией**. Благодаря наличию профага лизогенная бактерия обладает **иммунитетом** к заражению фаговыми частицами того же типа. Для поддержания иммунитета бактерии достаточно одного интегрированного профага, так что, как правило, бактериальный геном содержит лишь одну копию профага конкретного типа.

Бактериофаг может переходить из одной формы в другую: от лизогении к литическому циклу и обратно. На РИС. 14.1 показано, что фаг, образованный в литическом цикле, проникает в клетку бактерии-хозяина; он может либо повторить литический цикл, либо перейти в лизогенное состояние. Выбор пути зависит от условий инфекции и сочетания генотипов фага и бактерии.

Профаг может быть выведен из лизогенного состояния с помощью процесса, названного **индукцией**. Она проходит в несколько этапов. Сначала фаговая ДНК высвобождается из хромосомного окружения посредством **эксцизии** и только затем вступает в литический цикл.

Альтернативные формы размножения и распространения одних и тех же фагов поддерживаются разными системами контроля транскрипции. Лизогения обуславливается взаимодействием фагового репрессора с оператором. Литический цикл требует целого каскада реакций контроля транскрипции. Таким образом, переход из литического цикла в лизогению выполняется в результате репрессии, а обратный переход (индукция профага) происходит в результате отмены этой самой репрессии.

Другая стратегия выживания в бактериальной клетке представлена **плазмидами**. Плазмиды являются автономными единицами, существующими в виде независимых **внехромосомных геномов**. Плазмиды

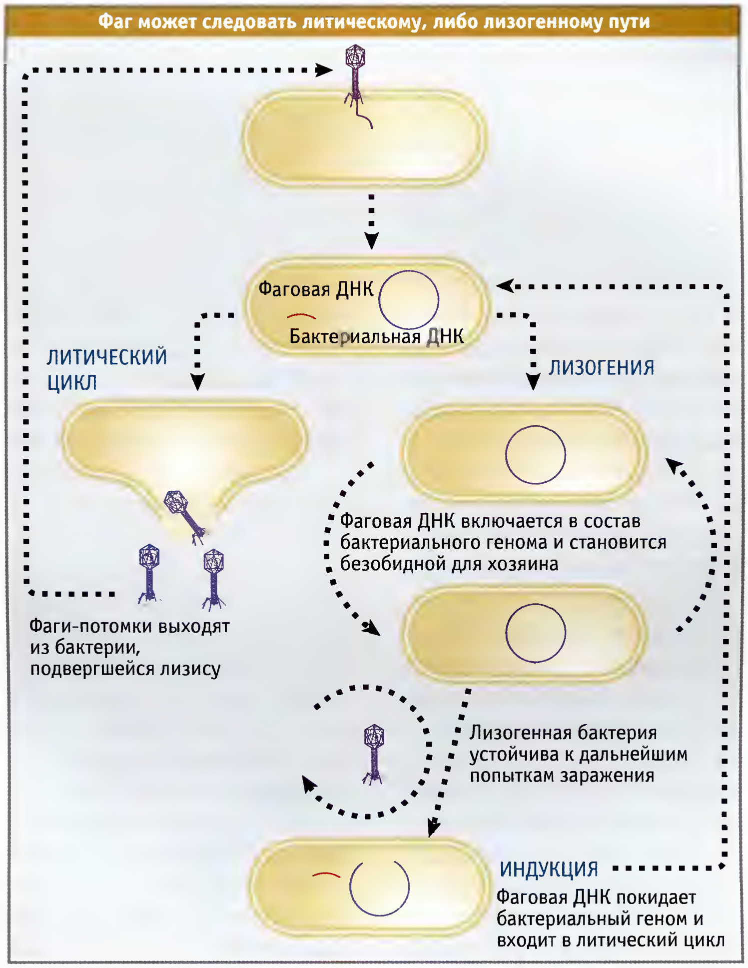


РИС. 14.1 Литический цикл бактериофага направлен на воспроизводство инфекционных частиц и приводит к гибели бактерии-хозяина. Лизогенная форма существования позволяет сохранять фаговый геном как часть генетической информации бактерии

Фаги и плазмиды населяют бактериальные клетки			
Тип генетической единицы	Структура генома	Способ распространения	Последствия для хозяина
Литический фаг	дц- или оцДНК или РНК линейная или кольцевая	Инфицирует хозяина	Гибель
Лизогенный фаг	дцДНК	Линейная область хромосомы хозяина	Иммунитет к инфекции
Плазмида	Кольцевая дцДНК	Реплицируется в заданном числе копий Может передаваться	Иммунитет к плазмидам той же группы
Эписома	Кольцевая дцДНК	Свободное кольцо либо линейная вставка	Распространение ДНК хозяина

РИС. 14.2 В бактериальных клетках встречается несколько типов независимых генетических единиц

можно охарактеризовать как самореплицирующиеся кольцевые молекулы ДНК, которые поддерживаются в характерном (стабильном) числе копий на клетку (то есть в ряду клеточных поколений это число не меняется).

У некоторых плазмид есть альтернативные способы существования. Взамен пребывания в автономном внехромосомном состоянии такая плазмида может быть превращена во вставку бактериальной хромосомы и вести себя как область этой хромосомы. Плазмиды, способные к таким превращениям, называют **эписомами**. В литературе есть тенденция к смешению понятий «плазмида» и «эписома»; два термина часто используются как взаимозаменяемые.

Стремление эгоистично обладать бактерией и не пускать в нее себе подобных роднит лизогенных фагов с плазмидами и эписомами. У плазмид эффект эгоистичного обладания также называют **иммунитетом**, хотя он имеет другую природу, чем у лизогенных фагов. (Контроль размножения плазмид мы обсудим в гл. 15 «Репликон».)

На РИС. 14.2 приведены все вышеперечисленные типы генетических единиц, распространяющих себя в популяциях бактерий в качестве независимых геномов. Геном литического фага может быть представлен нуклеиновой кислотой любого типа. Литические фаги осваивают популяцию бактерий через заражение инфекционными частицами все новых клеток. Геномы лизогенных фагов, плазмид и эписом находятся в форме двухцепочечной ДНК. Некоторые плазмиды и эписомы передаются между клетками в процессе конъюгации, которая протекает с установлением прямого контакта между клетками донора и реципиента. При любом из двух способов передачи генетического материала между клетками (как при заражении инфекционными частицами, так и при конъюгации) есть возможность передачи случайно захваченных бактериальных генов вместе с фаговой или плазмидной ДНК. Можно сказать, что наличие фагов и плазмид способствует возникновению потоков генетического материала в популяциях бактерий.

14.2 Литическое развитие можно разделить на две фазы

Ключевые понятия:

- Инфекционный цикл литического фага можно разделить на ранний период (события до репликации) и поздний период (после начала репликации ДНК фага).
- В ходе литического цикла единичный геном бактериофага воспроизводится во множестве копий, которые реплицируются и рекомбинируют.

Геном бактериофага невелик. Фаги — это вирусы, и размеры их геномов ограничены необходимостью укладки нуклеиновой кислоты в стандартный белковый капсид. Это ограничение является весомым фактором формирования стратегий размножения, используемых вирусами. Чаще всего вирус берет под свой контроль системы обеспечения функций хозяина и заставляет их заниматься экспрессией и репликацией вирусных генов — в ущерб нуждам самой клетки.

Большинство бактериофагов несет в себе гены, которые гарантируют их ДНК приоритет при репликации. Продукты этих генов имеют прямое отношение к инициации репликации; среди них может быть даже чужеродная для клетки ДНК-полимераза. Корректируются также клеточные системы обеспечения транскрипции (это может включать как замещение клеточной РНК-полимеразы на РНК-полимеразу фага, так и модификацию условий инициации и терминации транскрипции для клеточной РНК-полимеразы). Результат всегда один и тот же: продукция фаговых мРНК получает приоритет при транскрипции. Для синтеза белка фаги (большая их часть) ограничиваются аппаратом белкового синтеза хозяина, изменяя его активность путем замещения бактериальной РНК на фаговую.

Литическое развитие бактериофага представляет собой особый порядок, в котором его гены начинают свою экспрессию. Соблюдение этого порядка гарантирует присутствие нужного компонента в нужном количестве

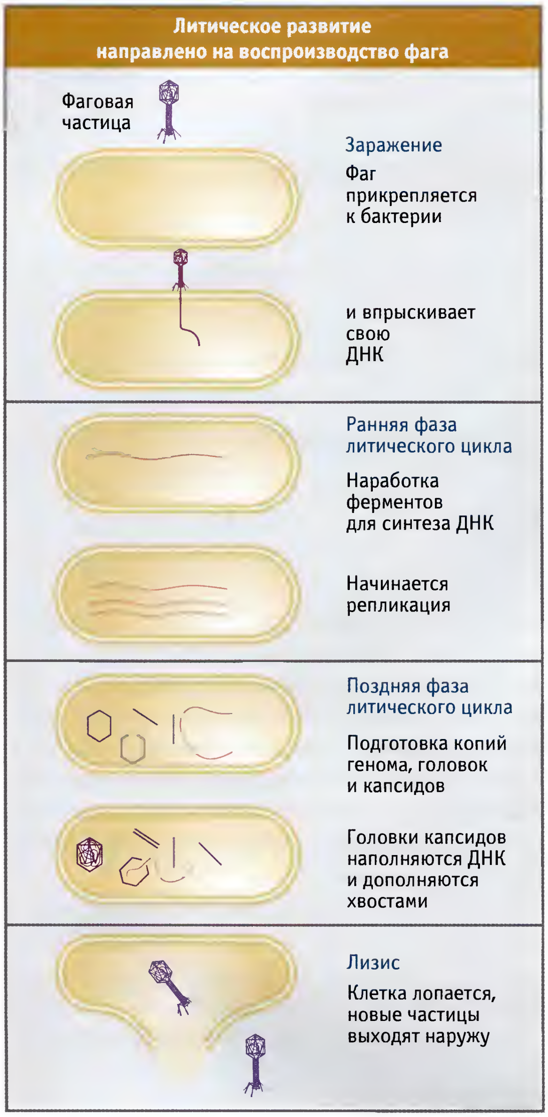


РИС. 14.3 Литическое развитие направлено на размножение геномных копий и белковых субъединиц для сборки частиц-потомков бактериофага

в нужное время. Как показано на РИС. 14.3, литический цикл можно разделить на два периода:

- **Ранняя инфекция** включает период от попадания в клетку чужеродной ДНК (или РНК) до начала ее репликации.
- **Поздняя инфекция** — это период от начала репликации фаговой ДНК до разрушения бактериальной клетки и выхода частиц-потомков фага.

Ранняя фаза инфекции отводится на наработку ферментов, которые в дальнейшем будут заняты воспроизводством ДНК. Это ферменты синтеза, рекомбинации, а иногда также ферменты модификации ДНК. В результате активности этих ферментов в клетке накапливаются (образуют *пул*) многочисленные копии фагового генома. Эти копии продолжают реплицироваться и рекомбинировать между собой, так что *события одного литического цикла, на самом деле, затрагивают целый пул фаговых геномов.*

В поздней фазе инфекции происходит синтез белковых компонентов будущих фаговых частиц. Для сборки частиц со сложной морфологией требуется множество различных белков, поэтому большая часть фагового генома кодирует поздние функции. В сборке фаговых частиц кроме структурных белков должны участвовать белки сборки фага (assembly proteins), которые помогают сборке, но в частицу не включаются. На фоне мероприятий по сборке головок и хвостов будущих частиц достигает пика своей скорости репликация фаговой ДНК. Наконец, пустые головки комплектуются копиями генома и хвостами. Клетка-хозяин подвергается лизису. Инфекционные частицы бактериофага выходят наружу, чтобы заражать другие клетки.

14.3 Литическое развитие подвержено каскадной регуляции

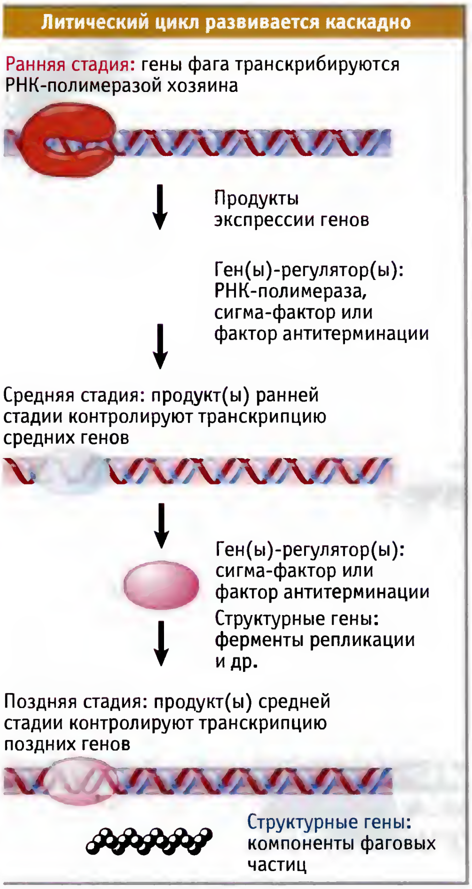
Ключевые понятия:

- Ранние гены бактериофага, транскрибируемые РНК-полимеразой хозяина сразу после заражения, служат регуляторами экспрессии средних генов.
- В группе средних генов есть регуляторы транскрипции поздних генов.
- Регуляция по типу каскада обеспечивает упорядоченную экспрессию дискретных групп генов в ходе литического развития.

У многих бактериофагов генетические карты отражают ход развития литического цикла. Концепция оперона достигает при этом своего крайнего выражения, поскольку у таких фагов гены, кодирующие белки со сцепленными функциями, собраны в кластеры, что делает контроль их экспрессии особенно экономичным и позволяет минимизировать число переключений в регуляторном контуре литического цикла.

Литическое развитие регулируется по принципу положительного контроля (это означает, что каждая группа генов бактериофага должна получить особый стимулирующий сигнал). На РИС. 14.4 показано, что экспрессия регуляторных генов происходит **каскадно**. То есть продукты генов, экспрессирующихся на одной ступени цикла, служат, прямо или косвенно, факторами транскрипции генов следующей ступени.

Первая волна экспрессии обязательно опирается на аппарат транскрипции клетки-хозяина. Первично экспрессирующиеся гены, как правило, немногочисленны, а их промоторы идентичны промоторам генов хозяина. Названия, данные этой группе генов у разных фагов, отличаются. Чаще всего их называют **ранними генами** (early genes). У фага лямбда они получили название **немедленноранних генов** (immediate early genes). Эти



На молекулярном уровне развитие литического цикла
яет собой регуляторный каскад, в котором продукт гена
дии необходим для экспрессии генов следующей стадии

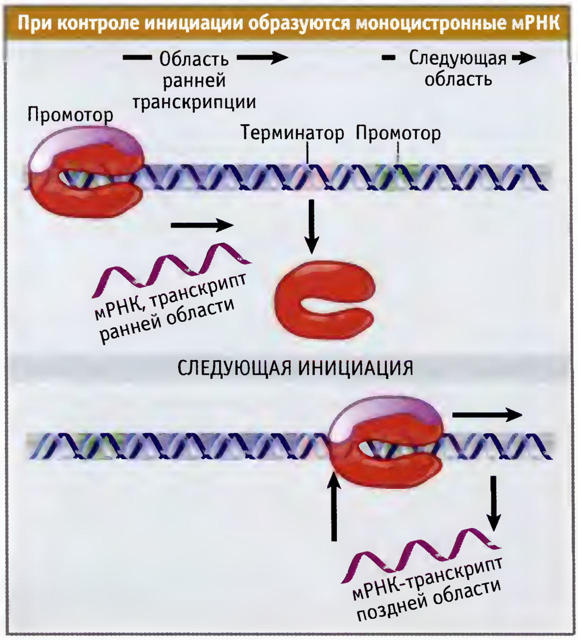


РИС. 14.5 Если контролю подлежит инициация транскрипции, «ранняя» и «поздняя» области генома бактериофага обслуживаются как независимые транскрипционные единицы, каждая со своим промотором и терминатором; с них считываются отдельные независимые мРНК. При контроле инициации ранние и поздние транскрипционные единицы не обязаны располагаться по соседству друг с другом

к бы их ни называли, служат распорядителями льной части раннего периода развития лити- цикла. В большинстве случаев их следует расце- сключительно как средство перехода на следую- іень каскада, и *один из этих генов обязательно п белок, необходимый для транскрипции генов ей группы.*

ая группа генов (**средние гены**, middle genes) вестна как **задержанноранние** (у фага лямб- ed early genes). Их экспрессия начинает детек- ся одновременно с появлением белка-регулято- одукта раннего гена (или генов). В соответствии еристиками конкретного регуляторного кон- этой стадии развития немедленноранние гены одолжить либо перестать экспрессироваться. гуляции подлежит *инициация* транскрипции (задержанноранних) генов, то немедленноран- ы больше не нужны и могут быть выключены). Если контрольным пунктом служит тер- і транскрипции средних генов, ранние гены ают экспрессироваться (см. РИС. 14.6). В любом ровень экспрессии собственных генов хозяина нденцию снижаться. Две первые волны экспрес- ов бактериофага, вместе взятые, обеспечивают

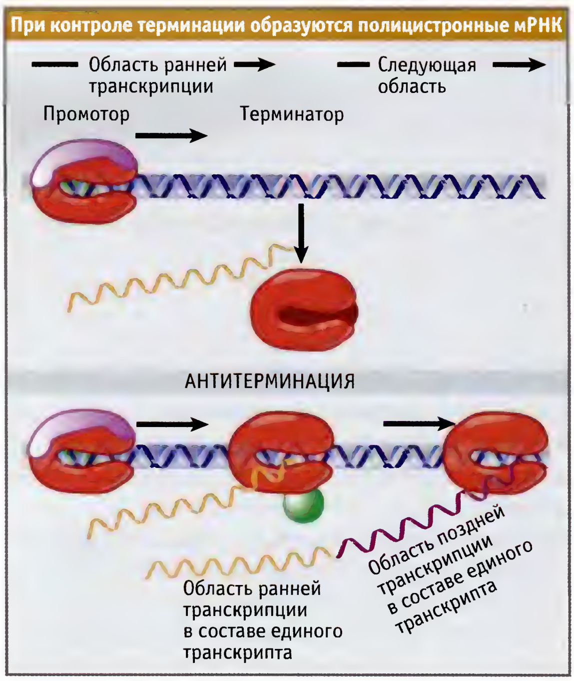


РИС. 14.6 Если контролю подлежит терминация транскрипции, ранние и поздние транскрипционные единицы должны располагаться стык в стык. После отмены терминации в середине литического цикла они станут одной транскрипционной единицей, с которой будет идти синтез полицистронных мРНК

выполнение всех функций, необходимых для его размножения (кроме сборки капсида фаговой частицы и лизиса бактериальной клетки).

Когда ДНК бактериофага начинает реплицироваться, наступает время **поздних генов** (late genes). Их транскрипция находится под контролем гена-регулятора из группы генов предыдущей, средней, стадии. Регулятором может быть еще один фактор антитерминации (у лямбда) или еще один альтернативный сигма-фактор (у SPO1).

Итак, литическое развитие на молекулярном уровне проходит в три стадии, как показано на рис. 14.4. Первые две стадии попадают в раннюю фазу инфекции; последняя стадия приходится на позднюю фазу инфекции (после начала репликации). Первая стадия знаменуется экспрессией ранних (немедленноранних) генов, транскрибируемых РНК-полимеразой хозяина при поддержке базового аппарата и факторов транскрипции того же хозяина. Продуктами первой стадии литического цикла во многих случаях являются одни лишь белки-регуляторы. На второй стадии идет транскрипция генов под контролем белков-регуляторов — эффекторов первой стадии. Большинство генов, транскрибируемых на второй стадии, кодирует ферменты репликации фаговой ДНК. В основе третьей, заключительной, стадии лежит экспрессия структурных генов и наработка компонентов капсида под контролем регулятора, синтезированного на второй стадии.

Ступенчатая система контроля, в которой каждая ступень является носителем регулятора для срабатывания следующей ступени, обслуживает литический каскад, в котором дискретные группы генов включаются (а иногда и выключаются) в определенный момент. Механистические принципы организации каскадов у разных фагов могут различаться, но результат срабатывания каскада всегда примерно один и тот же, что будет показано в следующих разделах.

14.4 Литический каскад контролируется двумя способами

Ключевые понятия:

- Белки-регуляторы, участвующие в литическом каскаде, либо поддерживают инициацию транскрипции с новых промоторов, либо принуждают РНК-полимеразу хозяина игнорировать кодоны-терминаторы.

На каждой стадии экспрессии генов проявляют активность гены-регуляторы, необходимость в продуктах которых возникает на следующей стадии развития цикла. У конкретного фага продукт гена-регулятора может быть либо РНК-полимеразой (новой для клетки), либо сигма-фактором (перепрофилирующим специфичность РНК-полимеразы хозяина; см. разд. 11.18 «Сигма-факторы организованы в каскады»), либо фактором антитерминации (который открывает РНК-полимеразе проход к новой группе генов; см. разд. 11.23 «Антитерминация — специальный регуляторный механизм»). Следующие два рисунка посвящены сравнению использования инициации и терминации транскрипции в качестве контрольных пунктов литического каскада.

Один из возможных путей литического каскада — это замена сигма-фактора бактерии своим собственным сигма-фактором. Это приводит к смещению профиля специфичности РНК-полимеразы к структуре промотора при инициации транскрипции (см. рис. 11.31), естественно, в пользу промоторов бактериофага.

Аналогичный подход — предложить клетке свою собственную РНК-полимеразу с высоким сродством к промоторам бактериофага — имеет такой же эффект: *подключение клеточных ресурсов транскрипции к новым промоторам.* На рис. 14.5 показано, что наборы транскриптов с разных ступеней каскада являются независимыми; поэтому после продукции нового сигма-фактора или РНК-полимеразы экспрессия ранних генов может прекратиться.

Альтернативным механизмом перевода литического каскада с ранней ступени на среднюю ступень может служить антитерминация. Возможность использования антитерминации обусловлена определенным порядком расположения генов в геноме бактериофага. На рис. 14.6 видно, что ранние гены лежат впритык к генам следующей очереди, но отделены от них сайтами терминации. После отмены терминации транскрипции в этих сайтах РНК-полимераза прочитывает их насквозь и, углубившись в следующий по ходу ген, прочитывает также и его (или группу генов). При использовании антитерминации нет нужды менять промоторы: новые гены экспрессируются в результате удлинения цепей РНК до молекул, у которых 5′-концевая область соответствует ранним генам, а 3′-концевая область — генам средней волны экспрессии. Неизбежным сопутствующим эффектом антитерминации является продленная экспрессия ранних генов.

У фага лямбда ген-регулятор *N*, контролирующий переключение литического каскада с немедленнораннней стадии на задержанноранннюю, идентифицируется по мутациям, приводящим к экспрессии *только* немедленноранних генов. Такие мутанты в своем инфекционном цикле не могут продвинуться дальше этой стадии (см. рис. 11.53). Похожий эффект можно наблюдать у фага SPO1, когда его ген *28* из-за мутаций не может служить полноценной матрицей для продукции σ^{gp28} (см. рис. 11.40). С точки зрения генетики, эффект новых параметров инициации и действие антитерминации похожи. Оба механизма обеспечивают положительный контроль, предусматривающий производство продукта раннего гена как обязательное условие экспрессии следующей группы генов. В принципе для построения литического каскада в каждом отдельном случае может быть достаточно наличия сигма-факторов либо факторов антитерминации различной специфичности.

14.5 Образование функциональных кластеров в геномах бактериофагов T7 и T4

Ключевые понятия:

- **Функционально связанные гены обычно образуют кластеры.**
- **Фаги T7 и T4 представляют собой пример регуляторных каскадов, в которых фаговая инфекция разделена на три периода.**

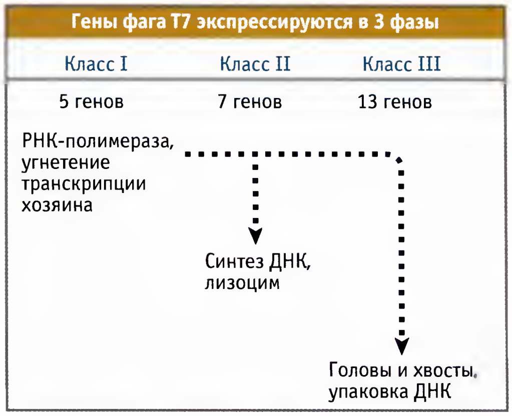


РИС. 14.7 У фага T7 три класса генов, экспрессирующихся последовательно. Размер генома составляет ~38 тпн

Геном фага T7 представлен тремя классами генов. Каждый из классов состоит из смежных локусов. Как показано на РИС. 14.7, гены класса I, немедленноранние, экспрессируются РНК-полимеразой хозяина, как только фаговая ДНК попадает в клетку. Среди продуктов экспрессии этих генов имеются ферменты, нарушающие экспрессию генов хозяина, и РНК-полимераза фага. РНК-полимераза фага ответственна за экспресиию генов класса II (которые в основном связаны с функцией синтеза ДНК) и генов класса III (участвуют в процессе сборки зрелой фаговой частицы).

Один из самых крупных фаговых геномов (165 Кб), геном фага T4, организован путем пространственной функциональной группировки генов. На РИС. 14.8 представлена генетическая карта фага T4. Гены, пронумерованные на рисунке, имеют *жизненно важное* значение (основные, или существенные гены): мутация в любом из этих локусов препятствует успешному завершению литического цикла. *Несущественные* гены отмечены трехбуквенными аббревиатурами. (Они считаются *второстепенными* при обычных условиях инфекции. Пока не совсем ясна роль многих второстепенных генов, но, предположительно, они могут предоставлять конкурентное преимущество в некоторых из сред обитания фага T4. В фаговых геномах меньшего размера большинство или все гены являются существенными.)

Гены T4 экспрессируются в три стадии. Краткое описание функций генов, работающих на каждой из стадий, дано на РИС. 14.9. Ранние гены транскрибируются РНК-полимеразой хозяина. Средние гены тоже транскрибируются РНК-полимеразой хозяина, но дополнительно необходимы два фаговых продукта — MotA и AsiA. В средних промоторах отсутствует консенсусная –30



РИС. 14.8 Генетическая карта T4 является кольцевой. Наблюдается пространственная кластеризация генов, кодирующих компоненты фага и важные процессы, такие как репликация ДНК. Но присутствует и рассеянное расположение генов, кодирующих разнообразные ферментативные и другие функции. Основные гены отмечены цифрами, второстепенные — буквами. На карте показаны только некоторые типичные гены T4



РИС. 14.9 **Литический каскад фага Т4 разделяют на две части: ранние функции связаны с синтезом ДНК и экспрессией генов; поздние функции связаны со сборкой фаговой частицы**

последовательность, а вместо нее присутствует последовательность, с которой связывается MotA. Белок фага является активатором, компенсирующим эту недостаточность промотора, и помогает РНК-полимеразе хозяина связаться с ним. (Этот механизм схож с механизмом, используемым фагом лямбда и проиллюстрированном на рис. 14.30.) Ранние и средние гены отвечают практически за все функции, связанные с синтезом ДНК фага, модификацией клеточных структур, транскрипцией и трансляцией фаговых генов.

Два основных гена транскрипционного типа выполняют регуляторную функцию: их продукты необходимы для экспрессии поздних генов. Интенсивность инфицирования фагом зависит от соотношения между репликацией ДНК и экспрессией поздних генов. Только активно реплицирующая ДНК может быть матрицей для транскрипции поздних генов. Связь создается новым сигма-фактором и другими изменениями РНК-полимеразы хозяина, делающими ее активной только на матрице реплицирующейся ДНК. Эта связь устанавливает корреляцию между синтезом белковых компонентов фага и количеством геномов, готовых для упаковки.

14.6 Немедленноранние и задержанноранние гены фага λ необходимы для лизогении и литического цикла

Ключевые понятия:

- **У фага лямбда два немедленноранних гена: *N* и *cro*. Они транскрибируются РНК-полимеразой хозяина.**
- ***N* необходим для экспрессии задержанноранних генов.**
- **Три из задержанноранних генов являются регуляторами.**
- **Для лизогении требуются задержанноранние гены *cII-cIII*.**
- **Для литического цикла требуется немедленноранний ген *cro* и задержанноранний ген *Q*.**

Один из самых сложных регуляторных каскадов используется фагом лямбда. Каскад литического развития фактически прямой, с двумя регуляторами, контролирующими успешность стадий развития. Схема регуляции литического цикла, тем не менее, смыкается со схемой запуска лизогении, как показано на РИС. 14.10.

Когда ДНК бактериофага лямбда попадает в клетку нового хозяина, сигнальные пути лизогении и литического цикла начинаются одинаково. Оба требуют экспрессии немедленноранних и задержанноранних генов, но затем появляется различие: литическое развитие происходит, если экспрессируются поздние гены, а лизогения — если синтезируется репрессор.

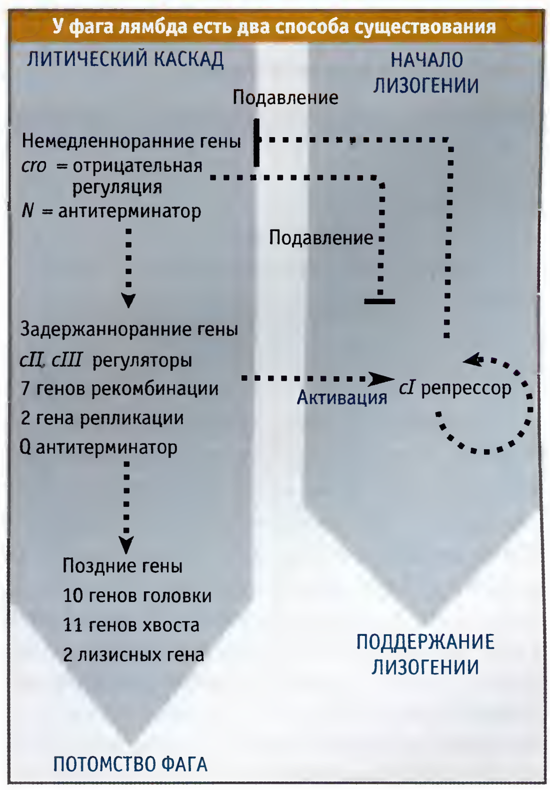


РИС. 14.10 **Литический каскад фага лямбда связан со схемой регуляции лизогении**

У фага лямбда лишь два немедленноранних гена, независимо синтезируемых РНК-полимеразой хозяина:

- *N* кодирует антитерминационный фактор, действие которого заключается в том, что на *nut*-сайте он позволяет перейти к транскрипции задержанноранних генов (см. разд. 11.24 «Антитерминация опосредована сайтами, олтичными от терминаторов).
- cro имеет двойную функцию: он предотвращает синтез репрессора (необходимое действие для начала литического цикла) и прекращает экспрессию немедленноранних генов (которые не нужны далее в литическом цикле).

Задержанноранние гены включают два гена для репликации (необходимые для литической инфекции), семь генов для рекомбинации (некоторые нужны для рекомбинации во время литической инфекции, а два необходимы для встраивания ДНК фага лямбда в бактериальный геном при лизогении) и три регулятора. Регуляторы имеют противоположные функции:

- Пара *cII-cIII* необходима для запуска синтеза репрессора.
- Регулятор *Q* является антитерминационным фактором, позволяющим РНК-полимеразе хозяина транскрибировать поздние гены.

Таким образом, задержанноранние гены служат двум целям: некоторые нужны для начала лизогении, а другие — для контроля литического цикла.

14.7 Литический цикл зависит от антитерминации

Ключевые понятия:

- **pN — фактор антитерминации, позволяющий РНК-полимеразе продолжать транскрипцию за границу двух немедленноранних генов.**
- **pQ — продукт задержанноprobраннего гена, антитерминатор, позволяющий РНК-полимеразе производить транскрипцию поздних генов.**
- **ДНК фага лямбда циркулирует после инфицирования; в результате, поздние гены образуют единственную транскрипционную единицу.**

Чтобы разобраться с двумя обозначенными выше путями транскрипции, давайте рассмотрим сначала литический цикл. РИС. 14.11 представляет схему ДНК фага лямбда. Группа генов, участвующих в регуляции, окружена генами, необходимыми для рекомбинации и репликации. Гены, кодирующие структурные элементы фага, сгруппированы. Все эти гены необходимы в литическом цикле и экспрессируются с трех промотеров с созданием полицистронных (кодирующих последовательности нескольких белков) РНК.

РИС показывает, что два немедленноранних ранних гена, *N* и *cro*, транскрибируются РНК-полимеразой бактерии-хозяина. *N* транскрибируется влево, а *cro* — вправо. Оба этих процесса транскрипции завершаются на конце гена. pN — регулятор, позволяющий транскрипции продолжаться в область задержанноранних генов. Он является фактором антитерминации, подавляющим действие терминаторов t_L и t_R (см. разд. 11.25, «Факторы терминации и антитерминации взаимодействуют с РНК-полимеразой»). В присутствии pN транскрипция продолжается влево от *N* на гены рекомбинации и вправо от *cro* на гены репликации.

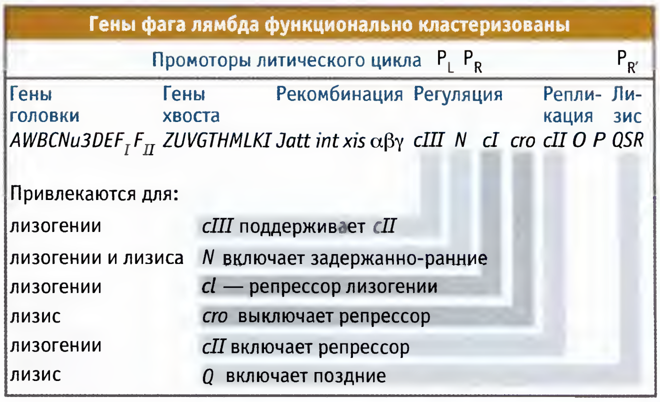


РИС. 14.11 Схема расположения групп генов с близкими функциями. Геном состоит из 48 514 пн

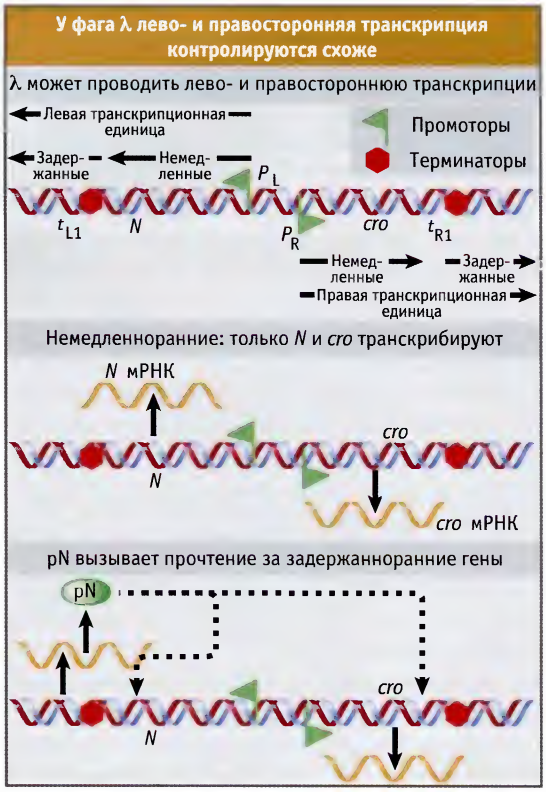


РИС. 14.12 Фаг лямбда имеет две ранних транскрипционных единицы. В левонаправленной единице верхняя цепь транскрибируется влево; в правонаправленной единице нижняя цепь транскрибируется вправо. Гены *N* и *cro* — немедленноранние гены. Они разделены задержанноранними генами — терминаторами. Синтез белка N позволяет РНК-полимеразе проходить терминатор t_{L1} влево и t_{R2} — вправо

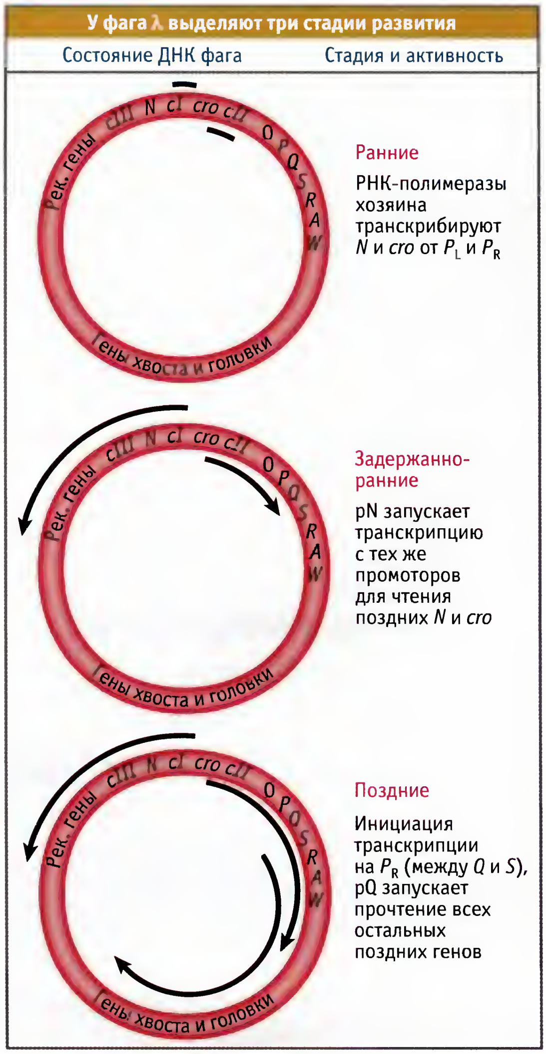


РИС. 14.13 **ДНК фага лямбда циклизируется во время инфицирования таким образом, что группа поздних генов соединяется в одну транскрипционную единицу**

Схема на рис. 14.11 показывает организацию ДНК фага лямбда в виде, в котором она существует в капсиде. Тем не менее через короткое время после заражения концы ДНК соединяются, образуя форму кольца. РИС. 14.13 показывает реальное состояние ДНК лямбда в течение заражения. Поздние гены объединены в отдельную группу, которая включает гены лизиса *S-R* с правого конца линейной ДНК, гены головки и хвостов *A-J* — с левого.

Поздние гены экспрессируются как отдельная транскрипционная единица, начинающаяся с промотора P_R, расположенного между Q и S. Поздний промотор используется всегда. В случае недостатка продукта гена Q (являющийся последним геном из правонаправленной задержанноранней транскрипционной единицы) транскрипция поздних генов завершается на сайте t_{R3}. Образующийся вследствие этого транскрипт составляет 194 основания и известен также как 6S РНК. Когда появляется pQ, он подавляет терминацию на t_{R3}, и синтез 6S РНК продолжается за терминатор, при этом экспрессируются поздние гены.

Транскрипция поздних генов не прекращается в каком-либо определенном месте, а продолжается на протяжении всех поздних генов до следующего участка. Аналогичное событие происходит с левонаправленной транскрипцией задержанноранних генов, обеспечивая функции рекомбинации. Транцкрипция в каждой части, вероятно, терминируется до того, как полимеразы могут столкнуться друг с другом.

14.8 Лизогения поддерживается белком-репрессором

Ключевые понятия:

- **Мутанты по гену *cI* не могут поддерживать лизогению.**
- ***cI* кодирует белок-репрессор, взаимодействующий с операторами O_L и O_R и блокирующий транскрипцию немедленноранних генов.**
- **Немедленноранние гены запускают регуляторный каскад; в результате их репрессия предотвращает развитие литического цикла.**

Изучая литический каскад фага лямбда, мы видим, что работа механизма транскрипции начинается на двух промоторах P_L и P_R для немедленноранних генов *N* и *cro*. Лямбда использует антитерминацию для продолжения экспрессии на следующей ступени (**задержанноранние** гены). Поэтому два одних и тех же промотора используются на протяжении всего раннего периода.

Подробная схема регуляторного участка изображена на РИС. 14.14. На нем показаны промоторы P_L и P_R, лежащие по бокам от гена *cL*. Каждый промотор включает оператор (O_L и O_R), с которыми связывается белок-репрессор, не позволяющий РНК-полимеразе начать транскрипцию. Последовательность каждого оператора перекрывается с последовательностью промотора, который он контролирует, поэтому эти последовательности часто объединяют в контролирующие участки P_L/O_L и P_R/O_R.

Так как механизм литического каскада имеет последовательный алгоритм развития, контролирующий участок выступает в роли пропускных ворот, которые контролируют весь цикл. *Белок-репрессор не дает РНК-полимеразе доступ к промоторам, таким образом не допуская участия генома фага в литическом цикле.* Механизм действия репрессора тот же, что и у репрессоров бактериальных оперонов: он связывается с специфичным оператором.

Белок-репрессор кодируется геном *cI*. Мутанты по этому гену не могут использовать лизогению, всегда вступая в литический цикл. Характеристика белка-репрессора привела к пониманию механизма защиты лизогена против сверхинфицирования новыми геномами фага лямбда.

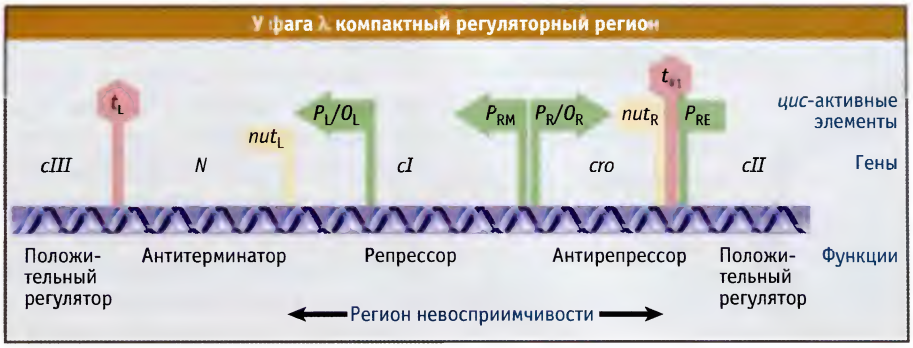


РИС. 14.14 Регуляторный участок генома фага λ содержит *цис*-действующие элементы и гены с *транс*-действующими функциями



РИС. 14.15 Бактерии, зараженные диким штаммом фага лямбда и вирулентными мутантами, могут быть различены по типу пятен. Воспроизведено по *Virology*, vol. 1, Kariser, D., pp. 423–443. Авторские права 1955, с разрешения Elsevier. Любезно предоставлено Dale Kaiser, Stanford University School of Medicine

Когда культура бактерий заражена фагом, на чашке Петри можно видеть небольшие участки, называемые **бляшками,** где клетки уничтожены вирусом. У дикого штамма фага бляшки мутные (затуманенные), потому что они содержат часть клеток, которые вместо лизинга перешли к лизогении. При мутации гена *cI* бляшки содержат только лизированные клетки, давая **прозрачные бляшки**. Три гена (*cI, cII* и *cIII*) были названы по их влиянию на фенотип. На РИС. 14.15 показаны бляшки, образующиеся у дикого и у мутантного штаммов.

Ген *cI* транскрибируется с промотора P_{RM}, который находится на его правом конце (*англ.* RM — repressor maintenance — поддержание репрессора). Транскрипция заканчивается на левом конце гена. мРНК начинается со старт-кодона AUG; из-за отсутствия сайта посадки рибосомы мРНК транслируется неэффективно, производя небольшое количество белка-репрессора.

14.9 Репрессор и его операторы формируют регион невосприимчивости

Ключевые понятия:

- Лямбдаподобные фаги обладают разными регионами невосприимчивости
- Лизогенный фаг делает клетку невосприимчивой к последующим заражениям фагами с таким же (как у него) регионом невосприимчивости.

Наличие репрессора объясняет феномен **невосприимчивости** (иммунности, immunity) клеток (и находящихся в них профагов) к повторному заражению. В случае если ДНК второго фага лямбда проникнет в лизогенную клетку, белок-репрессор, синтезированный с уже присутствующего в клетке профага, немедленно свяжется с операторами O_L и O_R новопроникшего генома и предотвратит вступление второго фага в литический цикл.

Исходно операторы были идентифицированы как сайты посадки репрессора при помощи так называемых **вирулентных** мутаций (*λvir*). Эти мутации предотвращают посадку репрессора на участки O_L и O_R, приводя к тому, что фаг сразу переходит к литическому пути в случае заражения новой клетки. Данные мутанты (*λvir*) способны расти и на уже зараженных фагом лизогенных бактериях, поскольку вирулентные мутации на участках O_L и O_R позволяют таким бактериофагам игнорировать присутствующий в клетке репрессор и запускать литический цикл. Вирулентные мутации фагов по своему эффекту эквивалентны мутациям операторов бактериальных оперонов.

Профаг переключается на литический путь в том случае, если лизогенный цикл будет прерван. Это происходит при инактивации репрессора (см. разд. 14.10 «ДНК-связывающая форма репрессора представляет собой димер»). Отсутствие репрессора позволяет РНК-полимеразе связаться с промоторами P_L и P_R, запуская

тем самым литический путь (как показано в нижней части рис. 14.26).

Эндогенная природа поддерживаемого репрессором цикла определяет быстрый ответ на любые изменения. Наличие репрессора необходимо для его собственного синтеза. Следовательно, экспрессия гена *cI* прекращается, как только присутствующий в клетке репрессор разрушен, а новые молекулы репрессора не могут быть наработаны, чтобы заменить поврежденные. Такой способ регуляции позволяет фагу переходить от лизогенного пути к литическому без конфликта двух этих стратегий.

Невосприимчивость клеток к повторному заражению фагом лямбда определяется регионом, включающим левый и правый операторы, ген *cI* и ген *cro*. Фаги с одинаковыми регионами (O_L,O_R, *cI* и *cro*) относятся к одной группе невосприимчивости, *поскольку данный регион определяет как сам репрессор, так и сайты его посадки.* Соответственно, данный регион принято называть **регионом невосприимчивости** или области иммунности (как показано на рис. 14.14). Каждый из четырех лямбдаподобных фагов (φ80, 21, 434 и λ) обладает собственным регионом невосприимчивости. Говоря о невосприимчивости клеток, содержащих лизогенный фаг, к повторному заражению фагом того же типа, точнее было бы сказать о невосприимчивости к любому фагу с таким же регионом невосприимчивости (независимо от степени различий в других частях генома).



РИС. 14.16 **В каждой субъединице репрессора N-концевая и C-концевая области полипептида организованы в отдельные домены. C-концевые домены двух субъединиц ассоциируют, давая димер. N-концевые домены связывают ДНК**

14.10 ДНК-связывающая форма репрессора представляет собой димер

Ключевые понятия:

- В мономере репрессора четко различимы два домена.
- ДНК-связывающий сайт находится в N-концевом домене репрессора.
- C-концевые домены репрессора димеризуются.
- Для связывания с оператором репрессору необходимо принять вид димера, чтобы два N-концевых домена могли одновременно контактировать с ДНК.
- Разобщение N- и C-концевого доменов (в результате протеолиза) снижает сродство репрессора к оператору и приводит к индукции литического цикла.

Субъединица (мономер) репрессора — это белок массой 27 кДа. Как показано на РИС. 14.16, в каждом мономере репрессора четко различимы два домена:

- N-концевой домен (аминокислоты 1-92) имеет сайт связывания с оператором.
- C-концевой домен (аминокислоты 132-236) отвечает за димеризацию.

Два домена соединены участком из сорока аминокислот. Протеолиз этого участка дает два полипептидных фрагмента, соответствующих разобщенным доменам.

Разобщенным доменам в принципе присуща некая активность. C-концевой фрагмент может образовывать олигомеры. N-концевой фрагмент может связываться с оператором, хотя показатели сродства (аффинность) к сайтам оператора у него гораздо ниже, чем у интактного репрессора. Можно сказать, что информация о специфичном контакте с ДНК заложена в самой структуре N-концевого домена; нельзя не отметить, что эффективность такого контакта многократно усилена C-концевым доменом.

Для поддержания лизогении необходимо присутствие димеров репрессора cI. Индукция лизогенного профага (побуждение к возвращению в литический цикл) происходит как следствие ограниченного протеолиза репрессора. (По большому счету, разрез в соединительном участке субъединицы лямбда репрессора, между аминокислотами 111 и 113, является аналогом аллостерической инактивации репрессора бактериального оперона. Только в данном случае репрессор лямбда, у которого нет свойства менять конформацию от связывания с низкомолекулярным индуктором, «обходится» протеолизом.) Индукция лизогенного профага обычно происходит при определенных (неблагоприятных) стечениях обстоятельств, например после ультрафиолетового облучения бактерий.

В интактном состоянии димеризация C-концевых доменов репрессора гарантирует одновременный контакт двух N-концевых доменов с ДНК. Протеолиз разобщает домены в субъединицах; как показано на РИС. 14.17, после протеолиза N-концевые домены разных субъединиц больше не способны димеризоваться. Кроме того что в общем пуле репрессора смещается равновесие между мономерами и димерами, протеолиз сводит на нет кооперативные взаимодействия между соседними димерами в операторах (см. разд. 14.16 «Кооперативный характер связывания повышает чувствительность регуляции»). В результате репрессор сходит с ДНК, что позволяет фагу сдвинуться с «мертвой точки» в развитии литического цикла.

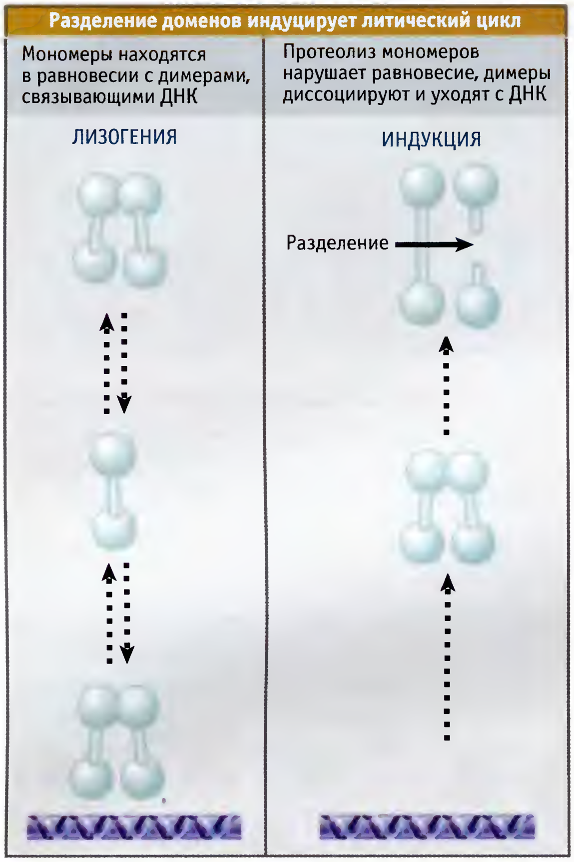


Рис. 14.17 **Димеры репрессора связываются с оператором. Сродство N-концевых доменов к ДНК зависит от димеризации С-концевых доменов**

Баланс между лизогенией и литическим циклом зависит от концентрации репрессора. Если для поддержания лизогении хватало определенной (небольшой) концентрации интактного репрессора, такой же концентрации обособленных N-концевых доменов будет уже *не*достаточно, потому что они будут иметь слишком низкое сродство к операторам. Очевидно, фаг должен обеспечить себе чрезвычайно высокую концентрацию репрессора, чтобы удержаться в лизогенном состоянии, несмотря на давление протеолиза.

14.11 Мотив «спираль-поворот-спираль» необходим репрессору для связывания ДНК

Ключевые понятия:

- Каждый ДНК-связывающий регион репрессора конъюгирует с полу-сайтом ДНК.
- ДНК-связывающий сайт содержит два коротких α-спиральных региона, которые ложатся в соседние витки большой бороздки ДНК.
- ДНК-связывающий сайт — это несовершенная палиндромная последовательно из 17 пн.

Итак, ДНК-связывающей единицей (модулем) репрессора является его димер. Он узнает в операторе последовательность длиной 17 пн, обладающую неполной симметрией относительно середины. Пример такого сайта приведен на Рис. 14.18. Последовательность каждой половины относительно центральной пары оснований иногда называют «полу-сайтом». С каждой половиной сайта контактирует один N-концевой домен репрессора. В непосредственном контакте с ДНК находятся две короткие α-спирали белковой глобулы. Подобный способ контакта характерен для целого ряда бактериальных ДНК-связывающих белков и даже для некоторых эукариотических факторов транскрипции (см. разд. 25.14 «Гомеодомены узнают в ДНК стандартную последовательность-мишень»).

В N-концевом домене репрессора лямбда есть несколько α-спиралей; они расположены примерно так, как показано на Рис. 14.19. Две из них отвечают за связывание ДНК. Следующий Рис. 14.20 иллюстрирует сам контакт по типу «спираль-поворот-спираль» (**«helix-turn-helix»**). Третья α-спираль каждого мономера репрессора состоит из девяти аминокислот и располагается под углом к семи аминокислотам второй α-спирали.

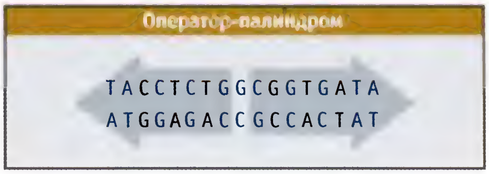


Рис. 14.18 **Оператор состоит из пары несовершенных инвертированных повторов по 8 пн, симметричных относительно центральной пары оснований (итого 17 пн). Полусайты отмечены голубым цветом, идентичные пары оснований — темно-синим цветом**



Рис. 14.19 **В N-концевом домене репрессора фага λ содержится пять структурных элементов типа «α-спираль». Вторая и третья спирали связывают ДНК**

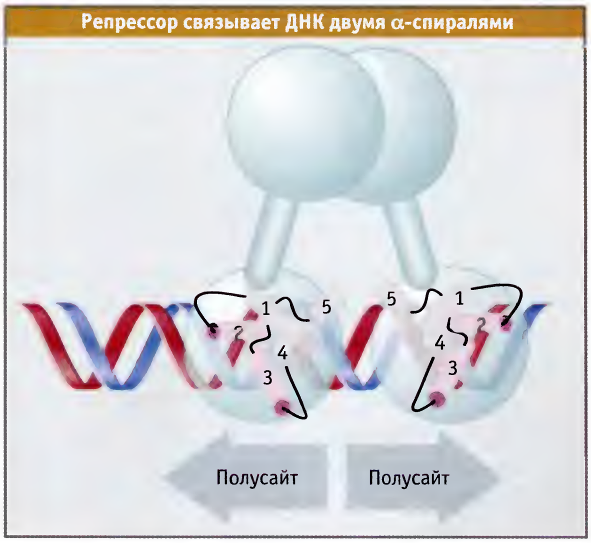


РИС. 14.20 Согласно «двуспиральной» модели связывания ДНК, третьи спирали двух субъединиц (мономеров) репрессора ложатся в большую бороздку по одну сторону двойной спирали ДНК. Вторые спирали лежат поперек бороздки

В димере репрессора третьи α-спирали двух мономеров разнесены на расстояние 34 Å (один виток двойной спирали); это позволяет им укладываться в последующие витки большой бороздки на одной стороне ДНК. (Вторые α-спирали двух мономеров, находясь под углом к третьим α-спиралям, должно быть, ложатся поперек большой бороздки ДНК.) Симметричный характер связывания означает, что два N-концевых домена (которые есть в каждом димере репрессора) имеют дело с похожими паттернами нуклеотидов на симметричных половинах сайта связывания в операторе.

14.12 «Спираль узнавания» задает специфичность к ДНК

Ключевые понятия:

- **Взаимодействие репрессора с конкретными парами нуклеотидов в операторе определяется аминокислотной последовательностью «спирали узнавания».**

Мотивы типа «спираль-поворот-спираль» репрессора лямбда широко распространены в бактериальных ДНК-связывающих белках: они есть в рецепторе цАМФ БРЦ, репрессоре *lac*, репрессорах некоторых других фагов. Способности этих белков к связыванию ДНК поддаются сравнению, и вот что можно сказать о роли каждой α-спирали:

- Взаимодействие третьей α-спирали с ДНК основывается на водородных связях между радикалами аминокислот и доступными атомами в парах азотистых оснований. Третья α-спираль ответственна за узнавание конкретной нуклеотидной последовательности-мишени, поэтому ее часто называют «спиралью узнавания» (recognition helix).
- Взаимодействие второй α-спирали с ДНК основывается на водородных связях с сахаро-фосфатным остовом двойной спирали. Это взаимодействие, хоть и необходимо для связывания, не влияет на специфичность узнавания мишени. В дополнение к этим контактам, значительная часть общей энергии связывания поступает за счет ионных взаимодействий репрессора с фосфатами ДНК.

Теперь немного поэкспериментируем: в репрессоре А третью α-спираль («спираль узнавания») заменим третьей α-спиралью от близкородственного репрессора В. (Такой «молекулярный конструктор» используется во многих лабораториях; объектом манипуляций служат нуклеотидные последовательности, кодирующие сегменты белка.) Получим новый белок (химеру). Его специфичность будет, как у В, донора третьей α-спирали. Таким образом, *аминокислотная последовательность короткого сегмента полипептида диктует всему белку его специфичность, органично вписываясь в контекст остальных сегментов полипептида.*

На рис. 14.21 изображены два белка фага лямбда, узнающие похожие нуклеотидные последовательности, а также детали их взаимодействия с ДНК. И у репрессора, и у белка Cro есть мотив «спираль-поворот-спираль»; тем не менее эти белки отличаются по специфичности к ДНК:

- Поддержание определенного положения α-спиралей друг относительно друга в том и другом белке

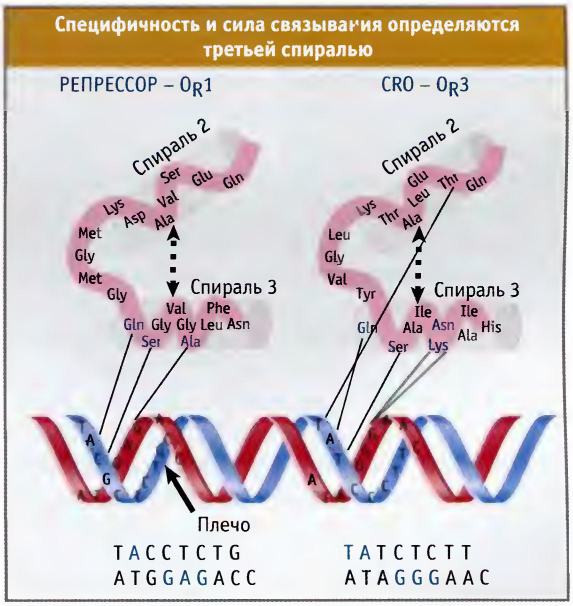


РИС. 14.21 Белки репрессор и Cro связывают ДНК двумя спиралями. Репрессор и Cro узнают операторы лямбда и удерживаются на них с эффективностью, зависящей от аминокислотной последовательности третьей спирали



РИС. 14.22 От основного массива репрессора (который находится по ту сторону ДНК) отходят N-концевые плечи; они «обнимают» ДНК (выходят на другую, ближнюю к нам сторону двойной спирали)

обеспечивается похожими взаимодействиями гидрофобных аминокислот: Ala-Val — у репрессора, Ala-Ile — у Cro.

- Аминокислоты в третьей спирали репрессора имеют контакты с конкретными парами азотистых оснований в операторе. Три аминокислоты репрессора узнают три основания в ДНК. У Cro аминокислоты в тех же (плюс дополнительных) положениях узнают пять (или, может быть, шесть) оснований в ДНК.

Две аминокислоты, участвующие в узнавании, у репрессора и Cro идентичны (это глутамин и серин в самом начале двойной спирали). Остальные контакты отличаются (аланин у репрессора против лицина и дополнительного аспарагина у Cro). Кроме того, у Cro треонин во второй спирали тоже непосредственно контактирует с ДНК.

На рис. 14.21 изображено связывание белков с последовательностями, которые подходят им наилучшим образом. В нижней части рисунка приведены другие варианты последовательностей с отличиями по трем из девяти нуклеотидных позиций. Использование похожих, но неидентичных, паттернов контакта между аминокислотами и азотистыми основаниями позволяет анализировать, каким образом сходство «спиралей узнавания» проецируется на сходство узнаваемых нуклеотидных последовательностей. По сути, репрессор и Cro узнают один и тот же набор последовательностей, но имеют разное сродство к конкретным последовательностям из этого набора.

Все азотистые основания, с которыми контактируют «спирали узнавания» репрессора и Cro, смотрят на одну и ту же сторону двойной спирали ДНК; это видно по рис. 14.21. Но у репрессора есть и дополнительный контакт на другой стороне двойной спирали. Если укоротить репрессор на первые шесть (N-концевых) аминокислот, предшествующих первой α-спирали, его сродство к ДНК уменьшится в 10^3 раз. Можно предположить, что первые шесть аминокислот отстоят от основного массива белковой глобулы (N-концевого домена) и образуют плечо, идущее полукольцом в обхват двойной спирали ДНК. На РИС. 14.22 показано, как лизины плеча взаимодействуют с гуанинами в большой бороздке, а также с фосфатным остовом ДНК.

Азотистые основания ДНК, не вступающие в прямой контакт с репрессором, могут, тем не менее, играть в его связывании очень важную роль. У фага *434* есть репрессор, похожий на репрессор у лямбда; он тоже связывает ДНК при помощи мотива «спираль-поворот-спираль». Анализ кристаллической структуры репрессора фага *434* показал, что третьи α-спирали контактируют с пятью дистальными (внешними) парами оснований сайта связывания и не контактируют с двумя проксимальными (внутренними) парами оснований. Однако сайты оператора с парами A-T в проксимальных положениях связывают репрессор крепче, чем сайты с парами G-C в тех же положениях. Объяснение этому следующее: связывание репрессора *434* слегка поворачивает ДНК в центре сайта для расширения угла между его половинами на ~3° (скорее всего, для достижения оптимальных условий связывания обеих субъединиц димера). Если в центре сайта стоит пара A-T, вращение идет легче; если в центре сайта стоит пара G-C — ДНК жестче сопротивляется вращению. Именно поэтому наличие в проксимальных положениях конкретных пар нуклеотидов влияет на сродство оператора к репрессору.

14.13 Димеры репрессора связываются с оператором кооперативно

Ключевые понятия:

- Связывание репрессора с одним из операторов повышает вероятность посадки второго репрессора на соседний оператор.
- Сродство репрессора к O_L1 и O_R1 примерно на порядок выше, чем к другим операторам, что определяет их первоочередное связывание.
- Кооперативность позволяет репрессору связывать сайты *01-02* при меньших концентрациях.

Каждый оператор содержит три сайта связывания репрессора. Как видно из РИС. 14.23, все шесть сайтов посадки репрессора разные, но все они удовлетворяют общей консенсусной последовательности. На каждом из операторов сайты связывания репрессора разделены A-T-богатыми спейсерами из 3-7 нуклеотидов. Сайты внутри каждого из операторов пронумерованы, давая O_L1-O_L2-O_L3 для левого оператора и O_R1-O_R2-O_R3 для правого оператора. В каждом из операторов сайт с номером 1 является ближайшим к точке начала транскрипции, а сайты с номерами 2 и 3 расположены выше по последовательности.

Как же репрессор определяет, с каким из трех сайтов он должен связаться? Дело в том, что и для правого, и для левого операторов первый сайт имеет примерно на порядок большее сродство к репрессору, чем второй и третий сайты. Следовательно, репрессор всегда сначала садится на сайты O_L1 и O_R1.

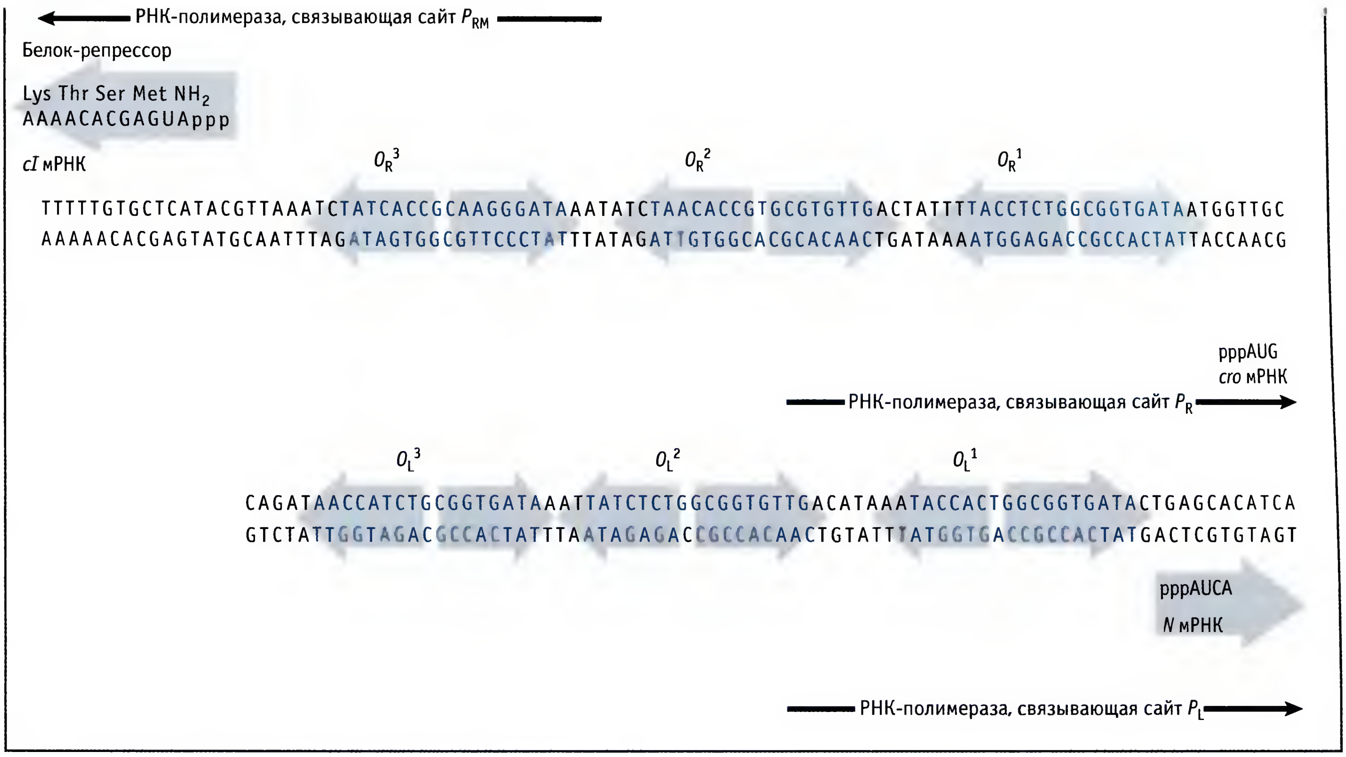


РИС. 14.23 Каждый оператор имеет 3 привязанных к рецептору сайта перекрывания промотора, с которым связана РНК-полимераза. Ориентация O_L развернута на 180 градусов для облегчения сравнения с O_R

Репрессор фага лямбда связывается с последующими сайтами на операторе по принципу кооперативности. Наличие димера на первом сайте существенно увеличивает сродство репрессора к сайту 2. Однако наличие репрессоров на сайтах 1 и 2 не влияет на возможность посадки репрессора на сайт 3. Таким образом, при стандартных лизогенных концентрациях репрессора сайты 1 и 2 заняты, а сайт 3 свободен.

В случае если сайт 1 неактивен (например, вследствие мутации), репрессор кооперативно связывается с сайтами 2 и 3. То есть посадка репрессора сразу на сайт 2 промотирует посадку другого димера на сайт 3. При этом существует прямое взаимодействие между димерами репрессора, не связанное с конформационными изменениями ДНК. С-концевой домен репрессора ответственен как за формирование димера, так и за кооперативное взаимодействие между димерами. РИС. 14.24 показывает, что каждая из субъединиц первого димера взаимодействует с аналогичной субъединицей второго димера, формируя тем самым тетрамер.

Целью кооперативного связывания является повышение сродства репрессора к оператору, что позволяет полностью связывать оператор при более низких концентрациях репрессора. Это крайне важно для системы, в которой снятие репрессии влечет необратимые последствия. Если для случая с бактериальным опероном несанкционированное снятие репрессии приведет лишь к нежелательному синтезу метаболических ферментов, снятие репрессии профага лямбда вызовет индукцию фага и лизис клетки.



РИС. 14.24 Когда два репрессора λ связаны с ДНК, каждая из субъединиц одного димера связывается с субъединицей другого димера

Из последовательностей, приведенных на рис. 14.23, видно, что O_L1 и O_R1 расположены более или менее в центре соответствующих сайтов посадки РНК-полимеразы P_L и P_R. Следовательно, посадка репрессора на O_L1-O_L2 и O_R1-O_R2 физически блокирует доступ РНК-полимеразы к соответствующим промоторам.

14.14 Находящийся на O_R2 репрессор взаимодействует с РНК-полимеразой на P_{RM}

Ключевые понятия:

- ДНК-связывающий регион репрессора на O_R2 взаимодействует с РНК-полимеразой и стабилизирует ее связывание с P_{RM}.
- Данный механизм является основой саморегуляции работы репрессора.

Для O_R и промотора P_{RM} показан целый спектр взаимодействий, определяющих транскрипцию *cI*. Сайт связывания РНК-полимеразы пристыкован к O_R2. Это объясняет то, как репрессор может аутогенно регулировать собственный синтез. Когда два димера репрессора связывают O_R1-O_R2, димер, находящийся на O_R2, взаимодействует с РНК-полимеразой (см. рис. 14.26 разд. 14.15 «Репрессор поддерживает цикл саморегуляции»). За данное взаимодействие отвечает N-концевой домен репрессора.

Мутации, отменяющие позитивный контроль, картированы в гене *cI*. В ходе экспериментов был обнаружен любопытный класс мутантов, у которых репрессор сохранял способность связывать оператор (подавляя транскрипцию), однако у них отсутствовала способность стимулировать РНК-полимеразу к началу транскрипции на промоторе P_{RM}. Мутации, приводящие к подобному эффекту, были картированы в регионе, кодирующем небольшую группу аминокислот, расположенных снаружи от второй спирали или в районе изгиба между второй и третьей спиралями. Мутации снижали отрицательный заряд данных участков молекулы репрессора. И наоборот, мутации, повышающие отрицательный заряд, обладали повышенной функцией активации РНК-полимеразы. Описанные эксперименты показали, что данная группа аминокислот формирует «кислый регион» на поверхности репрессора, взаимодействующий с основным регионом РНК-полимеразы за счет электростатических взаимодействий.



Рис. 14.25 Позитивно контролируемые мутации идентифицируют маленький район у спирали-2, которая непосредственно реагирует с РНК-полимеразой

Расположение таких «мутаций положительного контроля» в структуре репрессора показано на рис. 14.25. Они лежат в районе, расположенном вблизи фосфатных групп ДНК и одновременно рядом с РНК-полимеразой. Таким образом, группа аминокислот репрессора, вовлеченная в позитивную регуляцию, находится в непосредственном контакте с РНК-полимеразой. Здесь необходимо отметить, что *белок-белковое взаимодействие способно высвободить энергию, используемую для инициации транскрипции.*

Участок РНК-полимеразы, с которым контактирует репрессор, представляет собой σ70 субъединицу (принадлежащую фрагменту полимеразы, связывающему регион −35 промотора). Взаимодействие между репрессором и полимеразой необходимо для перехода полимеразы из состояния закрытого комплекса в состояние открытого комплекса (см. рис. 14.31).

Описанная схема объясняет, каким образом низкие концентрации репрессора могут усиливать его синтез. До тех пор, пока репрессора достаточно для связывания O_R2, РНК-полимераза будет продолжать транскрипцию гена *cI* с промотора P_{RM}.

14.15 Репрессор поддерживает цикл саморегуляции

Ключевые понятия:

- Связывание репрессора с O_L блокирует транскрипцию гена *N* с промотора P_L.
- Связывание репрессора с O_R блокирует транскрипцию гена *cro*, а также необходимо для транскрипции *cI*.
- Таким образом, связывание репрессора с операторами одновременно блокирует начало литического цикла и поддерживает его собственный синтез.

Репрессор связывается с двумя операторами независимо. Он имеет единственную функцию на O_L2, но обладает двумя активностями при связывании O_R. Эта особенность проиллюстрирована в верхней части рис. 14.26.

При связывании O_L репрессор оказывает влияние, аналогичное тому, что мы уже обсуждали для некоторых других систем: он не дает РНК-полимеразе начать транскрипцию на P_L, что, в свою очередь, останавливает экспрессию гена *N*. Поскольку промотор P_L используется фагом для транскрипции всех ранних генов левого региона, данное событие предотвращает экспрессию всего раннего транскрипта левого региона. *Таким образом, литический цикл оказывается блокирован еще на ранней стадии инфекционного цикла.*

Связывание репрессора с O_R1 предотвращает использование P_R. В результате *cro* и другие ранние гены правого региона не могут экспрессироваться (позже мы увидим, почему так важно исключить экспрессию *cro* в состоянии лизогении).

Присутствие репрессора на O_R имеет иные последствия. Промотор для синтеза репрессора, P_{RM}, вплотную

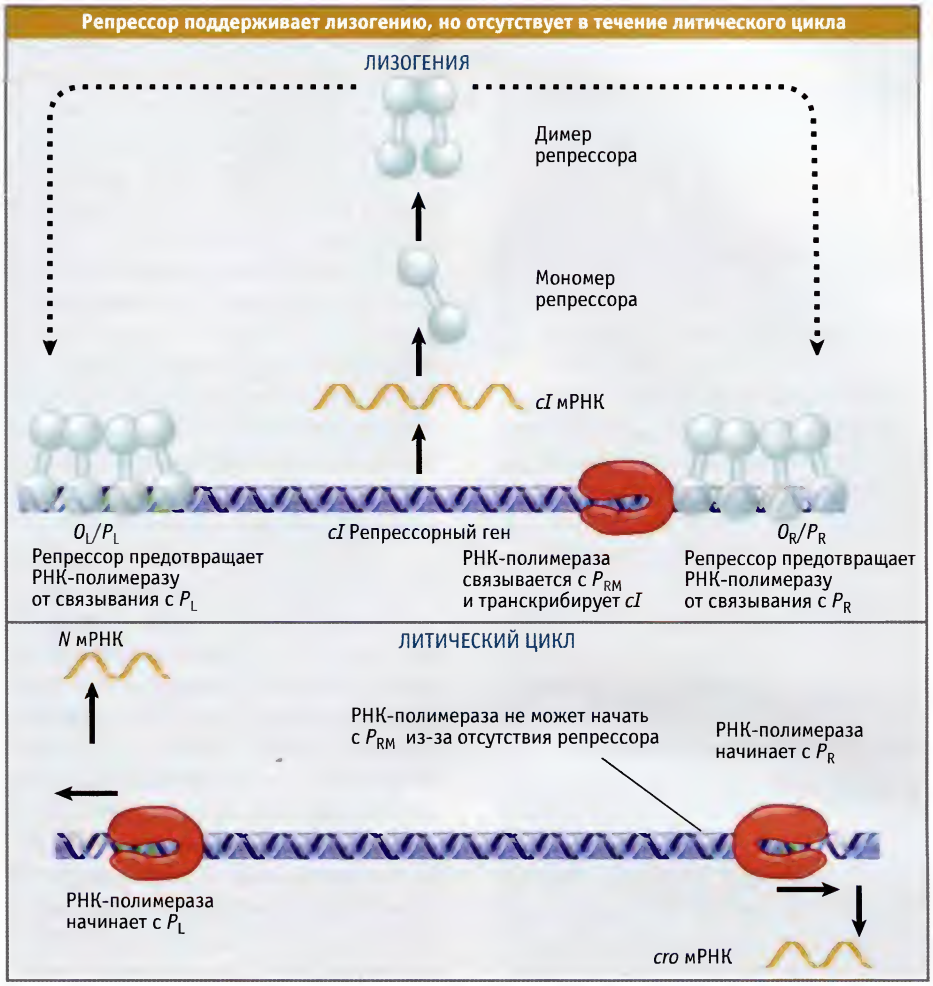


Рис. 14.26 Лизогения поддерживается циклом саморегуляции (верхняя часть). Если цикл прерывается, начинается литический цикл (нижняя часть)

пристыкован к оператору правого региона O_R. Такое расположение регуляторных участков ведет к тому, что РНК-полимераза может эффективно начинать синтез на промоторе P_{RM} лишь в том случае, если репрессор связан с оператором O_R. В данном случае репрессор выступает в качестве положительного регулятора, необходимого для транскрипции гена *cI* (см. разд. 14.14 «Находящийся на O_R2 репрессор взаимодействует с РНК-полимеразой на P_{RM}»). Поскольку репрессор является продуктом гена *cI*, данное взаимодействие создает цикл положительной обратной связи (саморегуляции), в котором наличие репрессора необходимо для его собственного синтеза.

Понимание природы данного регуляторного цикла объясняет биологическую природу лизогенного существования. Лизогения стабильна, поскольку описанный регуляторный цикл гарантирует, что до тех пор, пока концентрация репрессора находится выше определенного уровня, экспрессия гена *cI* продолжается. В результате O_L и O_R оказываются оккупированы репрессором неограниченно долго. За счет предотвращения перехода к литическому пути, описанный механизм поддерживает профаг в состоянии «спячки».

14.16 Кооперативный характер связывания повышает чувствительность регуляции

Ключевые понятия:

- Димеры репрессора, связавшие O_L1 и O_L2, взаимодействуют с димерами, связавшими O_R1 и O_R2; все четыре димера образуют октамер.
- Образование октамера приводит к сближению O_L3 с O_R3 и взаимодействию димеров репрессора.
- Кооперативный характер взаимодействия димеров репрессора с операторами повышает чувствительность регуляции.

В кооперативном связывании димеров репрессора участвуют оба оператора, левый и правый. Нормальное состояние операторов, занятых репрессором, — это удерживать димеры репрессора в первой и второй парах сайтов. Фактически на каждый оператор приходится тетрамер репрессора. Более того, две пары димеров образуют октамер (по сути, С-концевые

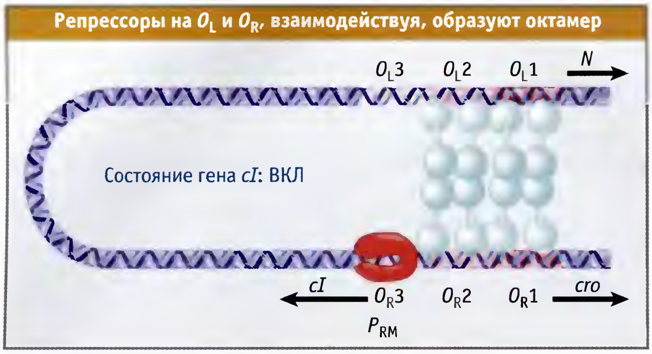


РИС. 14.27 В лизогенном состоянии субъединицы репрессора, связанные с O_L1 и O_L2, взаимодействуют с субъединицами, связанными с O_R1 и O_R2. При этом РНК-полимераза связывается с P_{RM} (который перекрывается с O_R3) и взаимодействует с репрессором, связанным с O_R2

домены восьми субъединиц кристаллизуются в виде октамера).

На РИС. 14.27 показано распределение димеров репрессора по сайтам операторов лизогенного профага. Димеры репрессора занимают сайты O_L1 и O_L2, O_R1 и O_R2. Димер, занимающий сайт O_R2, контактирует с РНК-полимеразой, начинающей транскрипцию с промотора P_{RM}.

Взаимодействие двух операторов (левого и правого, через репрессор) имеет особый смысл. Во-первых, оно стабилизирует посадку репрессора и тем самым сообщает операторам чувствительность к низким концентрациям репрессора. Во-вторых, посадка репрессора на O_R2 стабилизирует связывание РНК-полимеразы с P_{RM} для транскрипции *cI*. Таким образом, репрессор (продукт гена *cI*) даже в малых концентрациях *стимулирует продукцию самого себя.*

На участке между операторами O_L и O_R ДНК образует большую петлю (в ней содержится ген *cI*). Основание петли закреплено октамером репрессора. Октамер сближает третью (пока пустую) пару сайтов оператора, O_L3 и O_R3. В результате O_L3 и O_R3 также оказываются способны связать по одному димеру репрессора и взаимодействовать друг с другом, как показано на РИС. 14.28. Теперь, из-за занятости O_R3, РНК-полимераза не может использовать промотор P_{RM} для транскрипции *cI*: экспрессия репрессора выключается.

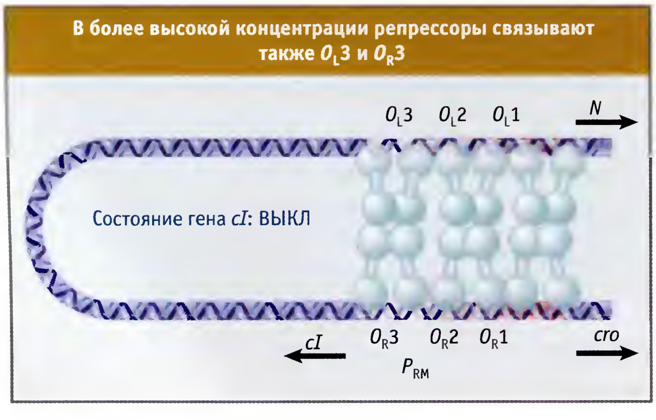


РИС. 14.28 В результате образования репрессором октамера, O_L3 и O_R3 оказываются сближены, поэтому увеличение концентрации репрессора позволяет еще одной паре его димеров связаться с O_L3 и O_R3 и взаимодействовать между собой

Эта модель объясняет чувствительность экспрессии *cI* к концентрации репрессора (его собственного продукта). В самом деле, в низких концентрациях репрессор образует на операторах октамер и активирует РНК-полимеразу (то есть регулирует собственную экспрессию по типу положительной обратной связи). Неизбежное увеличение концентрации репрессора приводит к связыванию им третьей пары сайтов, O_L3 и O_R3, и выключению транскрипции его собственного гена *cI* (регуляция по типу отрицательной обратной связи). У пороговых концентраций репрессора, требуемых для переключения между типами регуляции, очень низкие абсолютные значения. Система явно обладает повышенной чувствительностью; это объясняется кооперативным характером взаимодействия димеров репрессора с сериями сайтов в операторах. Незначительное изменение концентрации репрессора запускает соответствующий регуляторный ответ, направленный на стабилизацию лизогенного состояния.

Кооперативный характер взаимодействия репрессора с операторами снижает эффективную концентрацию репрессора (примерно в три раза). Соответствующее, относительно небольшое, количество репрессора должно быть уничтожено для выхода профага из лизогенного состояния. Поэтому индукция фага лямбда на молекулярном уровне отличается исключительно высокой эффективностью.

14.17 Гены *cII* и *cIII* нужны для установления лизогении

Ключевые понятия:

- Продукты задержанноранних генов *cII* и *cIII* необходимы, чтобы РНК-полимераза могла начать транскрипцию с промотора P_{RE}.
- Белок cII адаптирует промотор, а cIII защищает cII от деградации.
- Транскрипция с P_{RE} приводит к синтезу репрессора и мешает экспрессии *cro*.

В устройстве схемы поддержания лизогенного состояния лямбда есть один парадокс. *Присутствие белка-репрессора необходимо для его синтеза.* Этим объясняется, за счет чего может сохраняться состояние лизогении. Каков же процесс, запускающий синтез репрессора в первый раз?

Когда ДНК фага лямбда только внедряется в клетку нового хозяина (сразу после заражения), бактериальная РНК-полимераза не может транскрибировать *cI* из-за

отсутствия белка, который помогал бы РНК-полимеразе использовать промотор P_{RM}. В то же время отсутствие репрессора *cI* означает доступность промоторов P_R и P_L; с них немедленно начинается транскрипция генов *N* и *cro*. Вскоре после этого, при участии белка pN (антитерминатора) транскрипция может продолжаться далее. Удлинение транскрипта, идущего влево, захватывает область гена *cIII* (и еще нескольких генов; это равносильно активации их транскрипции). Транскрипт, идущий вправо, теперь содержит открытые рамки считывания *cII* и еще нескольких генов (см. рис. 14.14).

Сходство *cII* и *cIII* с *cI* состоит в том, что мутации в любом из них могут привести к появлению прозрачных бляшек. Однако мутанты по *cI* никогда не возвращаются в лизогенное состояние и не умеют его поддерживать. Мутанты по *cII* или по *cIII* испытывают трудности перехода в лизогенное состояние, но они способны поддерживать лизогению за счет цикла саморегуляции *cI*.

Мутанты ведут себя так, будто *cII* и *cIII* служат позитивными регуляторами перехода в лизогенное состояние. Действительно, продукты этих генов задействованы в *первичной* наработке репрессора cI с альтернативного промотора (не P_{RM}). Эта мера нужна только для того, чтобы *начать de novo* экспрессию *cI*, преодолевая принципиальную неспособность контура положительной обратной связи к самоинициации. Гены *cII* и *cIII* не предназначены для длительной опеки самоподдерживающегося контура экспрессии *cI*.

Белок cII напрямую регулирует экспрессию *cI* с промотора P_{RE}, который находится между генами *cro* и *cII*. (*англ.* RE, repressor establishment —установление репрессора). Этот промотор используется РНК-полимеразой только в присутствии cII, как изображено на РИС. 14.29.

В условиях *in vivo* белок cII чрезвычайно нестабилен: он очень быстро деградирует под действием бактериального белка HflA. Функция cIII сводится к защите cII от деградации.

Транскрипция с P_{RE} ведет к лизогении двумя путями. Прямой эффект транскрипции — это наработка репрессора (поскольку в транскрипт попадает вся область *cI* с открытой рамкой считывания). Косвенный эффект обусловлен тем, что транскрипция вначале проходит через ген *cro* в неверном направлении. Таким образом, 5'-концевая область транскрипта соответствует антисмысловой цепи *cro*. Гибридизуясь с аутентичной мРНК *cro*, транскрипт P_{RE} ингибирует ее трансляцию (в результате белок cro не синтезируется, литический цикл остается без позитивного регулятора; см. разд. 14.20 «Без *cro* полный литический цикл невозможен»).

По сравнению с «вялой» трансляцией транскрипта P_{RM}, кодирующая область *cI* в транскрипте P_{RE} транслируется очень эффективно. На самом деле, синтез репрессора с P_{RE} идет в семь-восемь раз интенсивнее, чем с P_{RM} по той простой причине, что в транскрипте P_{RE} есть эффективный сайт посадки рибосомы. В транскрипте P_{RM} сайт посадки рибосомы отсутствует (сразу идет инициирующий кодон AUG).

Для наработки репрессора используется специальный промотор

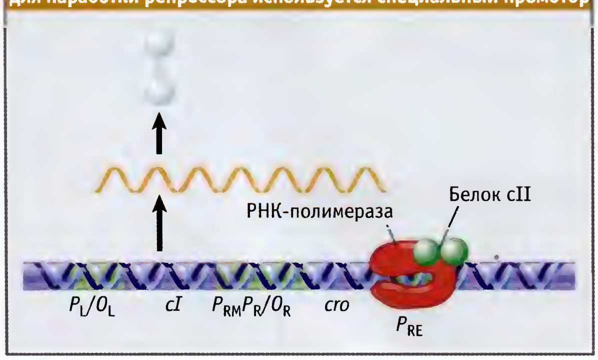


РИС. 14.29 Запуск наработки репрессора происходит совместными усилиями РНК-полимеразы и белка cII на промоторе P_{RE}. При этом инициируется транскрипция, идущая через *cro* к *cI* и далее

14.18 Слабому промотору нужен белок cII

Ключевые понятия:

- Промотор P_{RE} имеет нестандартные последовательности на функционально значимых участках −35 и −10.
- РНК-полимераза узнает этот промотор только в присутствии белка cII.
- Белок cII связывает участок −35 промотора P_{RE}.

Промотор P_{RE} плохо соответствует консенсусу для участка −10 и вообще не соответствует консенсусу для последовательности −35. Явная дефицитность этого промотора объясняет его функциональную зависимость от белка cII. Сам по себе P_{RE} не может быть транскрибирован РНК-полимеразой *in vitro*; для транскрипции необходимо добавить в систему очищенный cII. Адаптор cII связывает P_{RE} на участке от −25 до −45. Посадка РНК-полимеразы на адаптированный промотор закрывает участок от −12 до +13. Как следует из РИС. 14.30, адаптор и фермент взаимодействуют с близко расположенными (или даже перекрывающимися) сайтами промотора.

Важность нестандартных участков −35 и −10 для функции P_{RE} доказывается существованием группы мутаций *cy*. По своему эффекту они напоминают мутации в *cII* и *cIII*, мешающие переходу в лизогенное состояние; только *cy* являются *цис*-действующими. Мутации *cy* делят на две группы: *cyL* и *cyR*, находящиеся, соответственно, на участках −10 и −35 (то есть там, где в промоторе положено находиться консенсусам двух операторов).

Мутации *cyL* маркируют участок −10 и, скорее всего, мешают РНК-полимеразе узнавать промотор.

Мутации *cyR*, маркирующие участок −35, бывают двух типов: одни мешают посадке cII, другие —

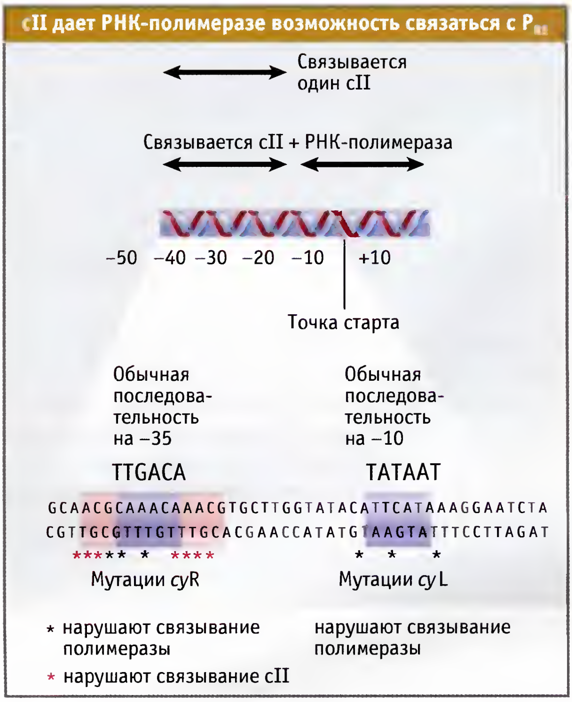


Рис. 14.30 РНК-полимераза связывается с P_{RE} только в присутствии cII, который связывается с P_{RE} в окрестности −35

посадке РНК-полимеразы. Мутации по центру участка не мешают связыванию cII; предполагают, что они мешают связыванию полимеразы. По обе стороны от них встречаются мутации, которые мешают связыванию cII; они приходятся на короткие повторы-тетрамеры TTGC. Каждая пара оснований в тетрамере удалена от гомологичной пары в другом тетрамере на один виток двойной спирали (10 пн). Так что, ложась на одну сторону ДНК, cII в принципе может узнавать сразу оба тетрамера.

cII играет роль позитивного регулятора; это означает, что посредник (accessory protein, белок доступа) cII повышает эффективность, с которой РНК-полимераза приступает к транскрипции. Рис. показывает, что регуляции может подлежать любой этап взаимодействия РНК-полимеразы с промотором P_{RE}. Помощь полимеразе может касаться как первичного связывания с образованием закрытого комплекса, так и последующего преобразования в открытый комплекс.

14.19 Установлению лизогении предшествует ряд событий

Ключевые понятия:

- cII и cIII оббеспечивают первичный всплеск синтеза репрессора, а также запускают ингибирование транскрипции поздних генов.
- С появлением репрессора транскрипция всех ранних генов выключается.
- Репрессор включает контур поддержания своего собственного синтеза.
- На заключительном этапе перехода в лизогенное состояние ДНК фага лямбда встраивается в геномную ДНК бактерии.

Теперь мы можем проследить за переходом в лизогенное состояние как возможным вариантом развития инфекции. Рис. восстанавливает картину ранних стадий развития инфекции и показывает, к чему приводит экспрессия генов *cII* и *cIII*. Наличие белка cII позволяет промотору P_{RE} использоваться для транскрипции области *cI*. С транскриптов P_{RE} быстро синтезируется большое количество белка-репрессора cI, который сразу же связывается с O_L и O_R.

Затем происходит прямое ингибирование транскрипции с P_L и P_R: связав оба промотора, репрессор выключает экспрессию всех генов бактериофага. Это останавливает синтез cII и cIII. Не отличаясь стабильностью, cII и cIII быстро сходят на нет. Без них экспрессия с P_{RE} полностью затухает, первичный синтез репрессора завершается.

Но, оказывается, на этот момент первичный синтез уже привел к установлению репрессора в качестве функциональной доминанты. Репрессор присутствует на O_R. Оттуда он запускает регуляторный контур своего собственного синтеза (работающий по принципу положительной обратной связи) с использованием промотора P_{RM}. Репрессор продолжает синтезироваться, хотя и с низкой интенсивностью.

Таким образом, установление лизогении начинается с синтеза репрессора на высоком уровне. Затем репрессор выключает все другие функции, при этом включается система поддержания его синтеза на низком уровне, достаточном для сохранения лизогенного состояния.

Положительная регуляция влияет на инициацию			
Промотор	Регулятор	Связывание промотора полимеразой (константа равновесия, K_B)	Конверсия полимеразы в открытую форму (константа скорости, k_2)
P_{RM}	репрессор	эффекта нет	11χ
P_{RE}	CII	100χ	100χ

Рис. 14.31 Положительная регуляция может поддерживать РНК-полимеразу на любом этапе инициации транскрипции

В этой главе мы не станем подробно останавливаться на других функциях, обуславливающих переход в лизогенное состояние (хотя такие функции есть). Например, инфицирующая ДНК фага лямбда должна встроиться в геном бактерии-хозяина (об этом см. разд. 19.17 «Сайт-специфичная рекомбинация сопровождается разрывом и воссоединением»). Для встраивания (внедрения, интеграции) необходим продукт гена *int*, экспрессирующегося с собственного промотора P_I, которому так же, как и P_{RE}, необходим посредник cII. Промоторы P_I и P_{RE} имеют похожие сайты связывания cII (но непохожие участки −10). Таким образом, функции, необходимые для установления лизогенного контролирующего контура, контролируются той же системой, что и функции, необходимые для интеграции ДНК фага в бактериальный геном. Можно предполагать, что единообразие регуляторных механизмов перехода фага в лизогенное состояние способствует большей четкости развития событий.

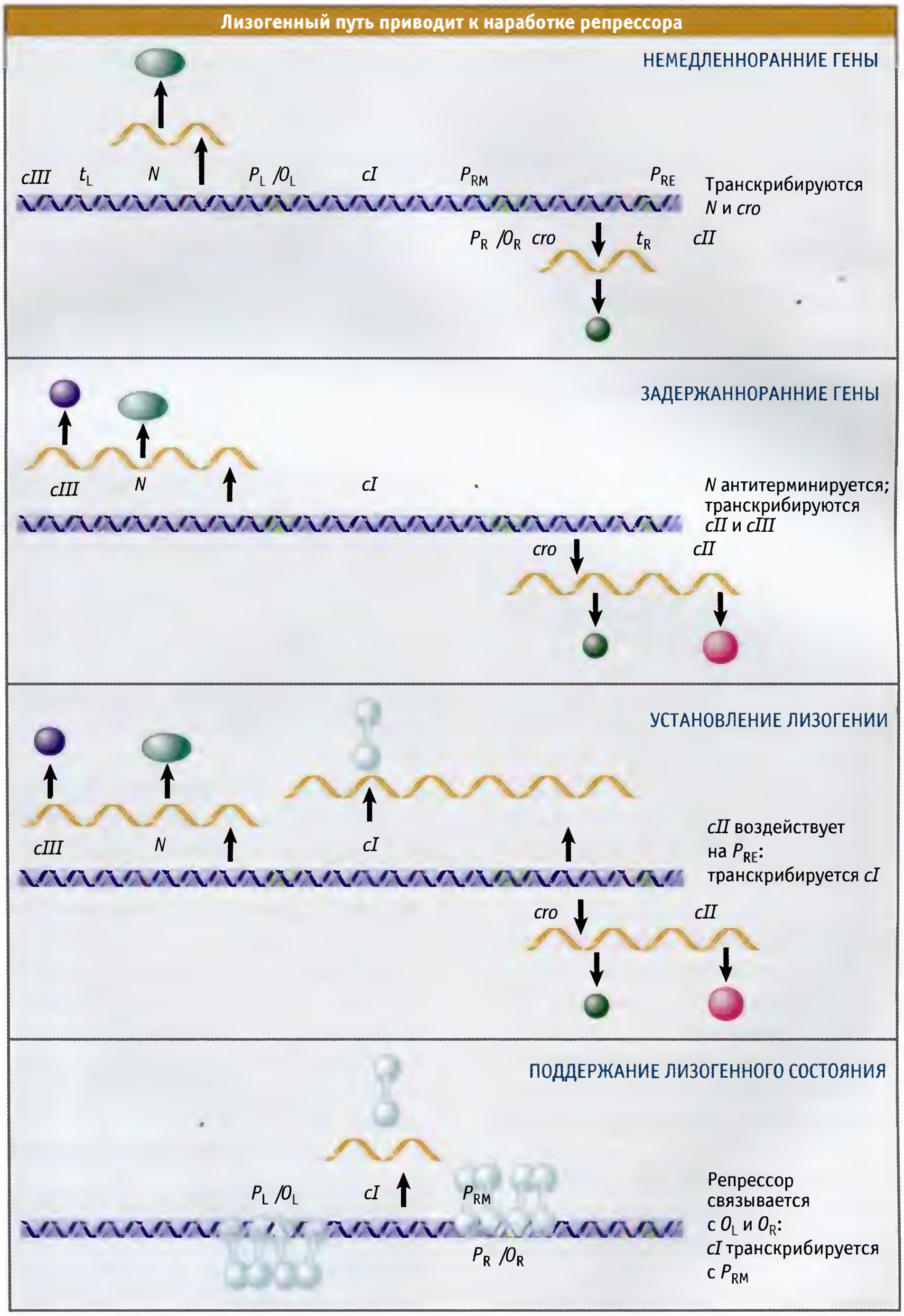


РИС. 14.32 Для установления лизогении необходим каскад, который вскоре выключается; на смену приходит регуляторный контур поддержания репрессора

Наконец, обратим внимание на то, что регуляторный каскад фага лямбда устроен весьма причудливо, с массой второстепенных контуров и перемычек. Известно, например, что белок cII способствует лизогении еще одним, косвенным, образом. Он поддерживает транскрипцию с промотора, который называется $P_{\text{anti-Q}}$; он находится внутри гена *Q*. Транскрипт $P_{\text{anti-Q}}$ — это антисмысловая версия гена *Q*; он гибридизуется с мРНК *Q* и препятствует ее трансляции (то есть синтезу белка Q, важнейшего регулятора поздней фазы литического цикла). Получается, что основной механизм перехода в лизогенное состояние имеет линейку «страховочных» приложений, способствующих достижению той же цели за счет ингибирования экспрессии генов *cro* и *Q* (позитивных регуляторов *завершения* литического цикла).

14.20 Без репрессора cro полный литический цикл невозможен

Ключевые понятия:

- Продукт гена *cro* связывается с теми же операторами, что и репрессор, но константы сродства у него другие.
- Если Cro связался с O_R3, он не дает РНК-полимеразе связаться с P_{RM} и, таким образом, блокирует самоподдержание репрессора.
- Если Cro связался с другими сайтами операторов O_R и O_L, он не дает РНК-полимеразе транскрибировать немедленно-ранние гены, что косвенным образом блокирует установление репрессора.

В начале инфекции у фага лямбда есть выбор: перейти в лизогенное состояние или развернуть инфекцию (литический цикл) по полной программе. Лизогения начинается с формирования регуляторного контура самоподдержания репрессора cI, после которого происходит глубокое ингибирование всего литического каскада.

Ключевой фигурой литического цикла выступает ген *cro*, кодирующий другой белок-репрессор. При нормальном развитии литического цикла *Cro отвечает за предотвращение синтеза репрессора cI*; полноценная активность Cro исключает возможность перехода в лизогенное состояние. Соответственно, фаги-мутанты по *cro* проявляют повышенную склонность к переходу в лизогенное состояние.

Белок Cro представляет собой небольшой димер (две субъединицы по 9 кДа), действующий внутри региона невосприимчивости. Это действие имеет два эффекта.

- Препятствует синтезу репрессора cI через контур самоподдержания, то есть препятствует транскрипции с P_{RM}.
- Препятствует экспрессии ранних генов с P_L и P_R.

Это можно интерпретировать так: в литическом цикле белок Cro должен вовремя выключить экспрессию ранних генов *и* не позволить бактериофагу перейти в лизогенное состояние.

Белок Cro выполняет свою функцию, связываясь с теми же операторами, что и репрессор cI. У Cro тоже есть мотив типа «спираль-поворот-спираль», в котором вторая спираль разнесена с третьей («спиралью узнавания») на довольно большой угол. (Этим, однако, сходство двух белков ограничивается; данный случай еще раз доказывает, что мотив «спираль-поворот-спираль» пригоден для использования в самых разных полипептидных контекстах.) Как и репрессор, Cro связывает операторы симметрично, двумя димерами на каждый сайт.

Аминокислотные последовательности мотивов «спираль-поворот-спираль» у Cro и cI похожи (так как они контактируют с одними и теми же сайтами ДНК; см. рис. 14.21). Контакты, которые Cro образует с ДНК, подобны контактам, образуемым cI, но *только с одной стороны*. В отличие от cI у Cro нет N-концевого домена, оборачивающегося вокруг двойной спирали.

Серия событий литического цикла воспроизведена на РИС. 14.33 (заметим, что в своей верхней части он идентичен рис. 14.32, посвященному лизогении). Как получилось, что Cro и cI, два похожих белка с одними и теми же сайтами-мишенями, являются антагонистами? Причина кроется в различной степени сродства (аффинности) двух белков к отдельным сайтам связывания внутри операторов.

Димеры Cro имеют большее сродство к O_R3, чем к O_R2 или к O_R1, вследствие чего Cro в первую очередь взаимодействует именно с O_R3. Это ингибирует связывание РНК-полимеразы с промотором P_{RM}. Таким образом, Cro приоритетно предупреждает включение контура поддержания лизогении.

Затем Cro связывается с O_R2 или O_R1. Аффинность белка к каждому из этих сайтов примерно одинакова (и без эффекта кооперации, как у репрессора). Присутствия Cro на O_R2 либо на O_R1 достаточно, чтобы РНК-полимераза больше не могла садиться на P_R. Из-за этого экспрессия ранних генов останавливается (в частности, прекращается продукция самого Cro). Вследствие нестабильности адаптора cII (продукта раннего гена, *cII*) использование промотора P_{RE} (первичный синтез репрессора) вскоре прекращается. Таким образом, первый и второй эффекты Cro *полностью* блокируют продукцию репрессора.

Что касается продолжения литического цикла, Cro подавляет экспрессию ранних генов, хотя и не полностью. Его неполный эффект объясняется довольно низким сродством Cro к O_L2 и O_R2 (оно в ~8 раз ниже, чем у репрессора cI). На самом деле, эффект Cro как репрессора ранних генов проявляется позже, в присутствии pQ. К этому времени начинается экспрессия поздних генов фага, и он сосредоточивается на продукции частиц-потомков.

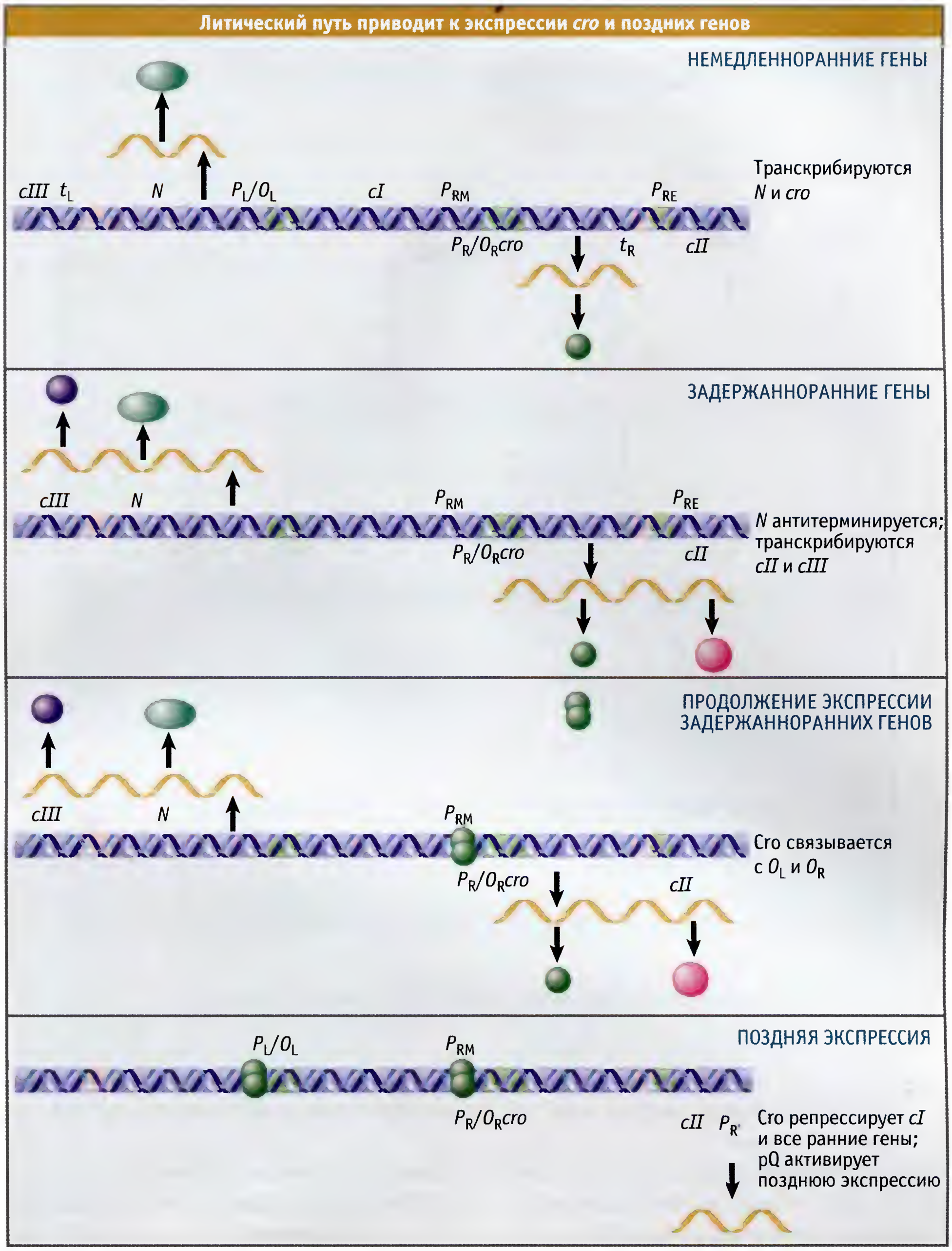


Рис. [illegible] В литическом каскаде белок Cro незаменим. Он напрямую препятствует поддержанию репрессора через P_{RM}, а также выключает экспрессию генов средней волны (задержанноранних) и тем самым мешает репрессору занимать O_L и O_R

14.21 На чем основывается баланс между лизогенией и литическим циклом?

Ключевые понятия:

- Лизогения и литический цикл развиваются по одному сценарию — вплоть до момента одновременного присутствия Cro и лизогенного репрессора cI.
- Ключевой момент: удастся ли белку cII стимулировать наработку репрессора в количестве, достаточном для продолжения эффекта действия Cro?

Программы развития событий литического цикла и перехода в лизогению настолько тесно связаны между собой, что судьба конкретного фага, поразившего нового хозяина, совершенно непредсказуема. Как завершится противостояние лизогенного репрессора и Cro? Установлением саморегулирующего контура, показанного на рис. 14.32, или развитием поздней экспрессии, как на рис. 14.33?

Обе программы развития начинаются одинаково: экспрессия немедленноранних генов перетекает в экспрессию задержанноранних. Непреодолимые различия между двумя программами возникают в момент приня-

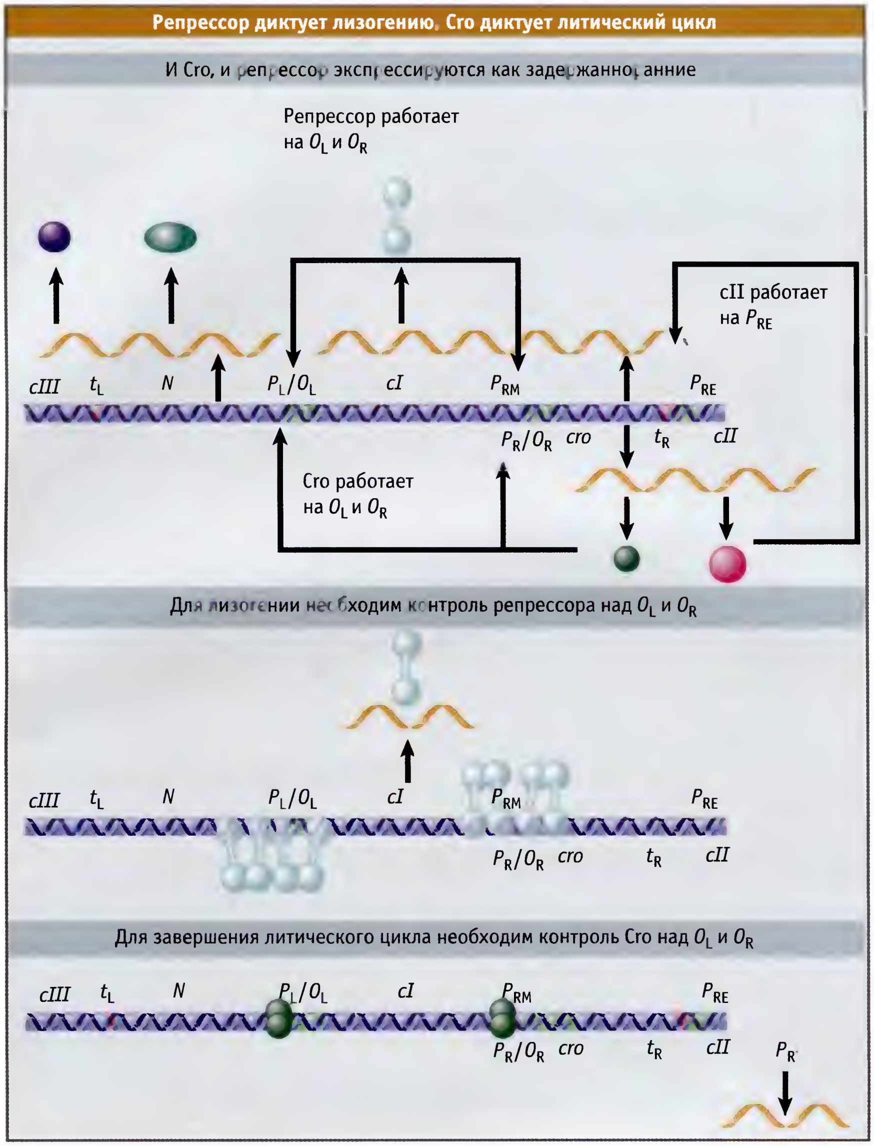


РИС. 14.34 Критический момент выбора между лизогенией и лизисом наступает после «задержки», во время экспрессии генов второй волны. Если cII удастся обеспечить синтез достаточного количества репрессора, это приведет к лизогении, поскольку операторы будут заняты репрессором. В противном случае с операторами будет взаимодействовать Cro, и тогда бактериофаг совершит полный литический цикл

тия решения о том, что присоединится к двум операторами — репрессор или белок Cro?

Продолжительность ранней фазы инфекции, в течение которой (так или иначе) принимается решение, всегда лимитирована. Независимо от того, какой путь будет выбран фагом, экспрессия всех ранних генов будет заблокирована (инактивацией промоторов P_R и P_L), и, вследствие исчезновения cII и cIII, первичная продукция репрессора (через промотор P_{RE}) прекратится.

Таким образом, выбор программы развития обусловлен событиями молекулярного уровня. Сумеет ли транскрипция с промотора P_{RE} поддержать активацию промотора P_{RM}? Если да, то переход в лизогенное состояние неизбежен. Если P_{RM} не активируется, то регулятор pQ приведет фага к литическому развитию. РИС. 14.34, последний в этой главе, посвящен критическому моменту совместного присутствия репрессора и Cro.

Начальный этап перехода в лизогенное состояние — это связывание димеров репрессора с O_L1 и O_R1. Взаимодействие быстро подхватывается кооперативными взаимодействиями еще двух димеров репрессора с O_L2 и O_R2, что ведет к полному выключению синтеза Cro и началу поддерживающего синтеза репрессора (через транскрипцию с P_{RM}).

Исходным событием для вступления в литический цикл является связывание белка Cro в операторе O_R3. Это блокирует систему сохранения лизогении на стадии инициации транскрипции в промоторе P_{RM}. Затем Cro должен

связаться с O_R1 или O_R2, а также с O_L1 или O_L2, чтобы подавить экспрессию ранних генов. В результате экспрессия генов *cII* и *cIII* будет приостановлена, что приведет к прекращению синтеза репрессора с промотора P_{RE}. Выключение процесса установления репрессии происходит, когда нестабильные белки cII и cIII распадаются.

cII является ключевым фактором в переходе от лизиса к лизогении. Когда cII активен, синтез репрессора через размещение на промоторе эффективен, в результате чего репрессор получает преимущество и на операторе. Когда cII не проявляет активности, место репрессора на операторе занимает *cro*. Количество белка cII при любых обстоятельствах определяет результат инфекции. Мутации, которые повышают стабильность белка cII, повышают и частоту переходов в лизогенное состояние. Такие мутации могут иметь место как в самом cII, так и в других генах. Вспомним, что причина нестабильности cII заключается в его уязвимости для протеаз хозяина. Поэтому уровень cII зависит не только от экспрессии стабилизатора (*cIII*), но и от функционирования клетки-хозяина. Эффект присутствия белка cIII, на самом деле, вторичен: cIII лишь защищает cII от деградации (тем более, что он сам нестабилен). Присутствие cIII не является гарантией выживания cII, однако в отсутствие cIII cII практически всегда инактивирован.

С другой стороны, в геноме бактерии-хозяина тоже могут быть мутации, повышающие частоту переходов в лизогенное состояние. Такие мутации имеют место в генах *hflA* и *hflB* (от *англ.* high frequency lysogenization — высокая частота лизогенизации). Эти мутации стабилизируют cII, поскольку инактивируют именно те протеазы хозяина, которые в норме занимаются протеолитической деградацией cII.

Получается, что на уровень cII (то есть на принятие решения о лизогенизации) влияет клетка-хозяин, которая, в свою очередь, находится под действием условий внешней среды. Так, протеазы, которые расщепляют cII, активируются при росте клеток в обогащенной среде. Можно сказать, что фаг лямбда склонен лизировать хорошо растущие бактериальные клетки. Лизогенное состояние фага лямбда, напротив, более типично для клеток, испытывающих трудности с пропитанием (фагу может просто не хватить ресурсов для эффективного литического цикла).

14.22 Резюме

Жизненный цикл бактериофага имеет своей целью продукцию большого числа инфекционных частиц. Этот цикл называют инфекционным (потому что он соответствует индивидуальному развитию инфекции у бактерии-хозяина), а также литическим циклом (потому что развитие инфекции завершается лизисом: гибелью бактерии-хозяина с разрывом клеточной стенки и выходом инфекционных частиц). У некоторых фагов возможна альтернативная, умеренная, форма существования, или лизогения. Геном лизогенного фага является частью генома бактерии-хозяина. ДНК лизогенного фага встроена в хромосомную ДНК бактерии; в такой инертной форме весь генетический материал фага наследуется как часть бактериальной хромосомы.

В типичном литическом цикле можно выделить три фазы. В первой фазе очень малое число генов фага транскрибируется РНК-полимеразой бактерии-хозяина. Один из этих генов (или два; редко больше) служит регулятором — он контролирует экспрессию группы генов во второй фазе инфекции. По тому же принципу продукты генов второй фазы регулируют экспрессию генов в третьей фазе инфекции.

В первых двух фазах литического цикла экспрессируются гены ферментов, необходимых для воспроизводства ДНК бактериофага; в третьей фазе экспрессируются гены, кодирующие структурные компоненты инфекционных частиц. Самым ранним генам литического цикла свойственно выключаться с наступлением второй и третьей фаз.

Гены фага лямбда организованы в группы, находящиеся под контролем обособленных регуляторных схем. Немедленноранний ген *N* кодирует белок-антитерминатор для транскрипции генов следующей волны экспрессии (задержанноранней; с промоторов P_R и P_L). Один из этих генов, *Q*, (вернее, его продукт, pQ) разрешает транскрипцию всех оставшихся (поздних) генов фага лямбда с промотора $P_{R'}$.

Развитие инфекции (ход литического цикла) может быть заблокировано за счет экспрессии гена *cI*. Его продукт, белок-репрессор cI, работает на операторах O_R и O_L, препятствуя транскрипции с промоторов P_R и P_L; кроме того, cI обеспечивает переход в лизогенное состояние и его стабилизацию. В геноме лизогенного фага экспрессируется один-единственный ген, *cI*, со своего промотора P_{RM}. Транскрипция с P_{RM} регулируется самим репрессором; регуляция идет по типу обратной связи с концентрацией репрессора.

Операторы лямбда O_R и O_L имеют по три сайта связывания репрессора. Каждый сайт — это палиндром, состоящий из двух симметричных половин. Репрессор cI функционирует в виде димера; каждая половина сайта связывания вступает в контакт с мономером репрессора.

N-концевой домен репрессора содержит мотив «спираль-поворот-спираль»; этот мотив контактирует с ДНК. Третья спираль в мотиве (так называемая спираль узнавания) ответственна за специфичность связывания (то есть контакт с конкретными нуклеотидными парами в операторе). Вторая спираль помогает третьей принимать правильное положение относительно ДНК; кроме этого, вторая спираль участвует в активирующем контакте с РНК-полимеразой на промоторе P_{RM}.

Противоположный, С-концевой, домен требуется репрессору для димеризации. Связывание димеров репрессора с оператором является кооперативным (это означает, что после того, как один димер связал первый сайт, связать соседний сайт второму димеру уже гораздо легче). Индукция профага лямбда (побуждение к выходу из лизогенного состояния и вхождению в литический цикл) вызывается ограниченным протеолизом репрессора (в результате отделения С-концевых доменов репрессор перестает функционировать в виде димеров; это

снижает сродство репрессора к ДНК, и лизогенное состояние больше не поддерживается).

Похожие мотивы «спираль-поворот-спираль» можно встретить и в других ДНК-связывающих белках, например, в белке Cro. Cro связывается с теми же операторами, что и репрессор, но имеет другие показатели сродства к конкретным сайтам (это обусловлено отличиями в аминокислотной последовательности третьей спирали). Cro связывает сайты оператора в индивидуальной манере (не кооперативно), начиная с O_R3. Cro накладывает запрет на переход в лизогенное состояние и обеспечивает нормальный ход литического цикла. Первоочередное связывание Cro с O_R3 препятствует синтезу репрессора с P_{RM}; последующие взаимодействия Cro с сайтами O_R2 и O_R1 (и сайтами O_L1 и O_L2) останавливают затянувшуюся экспрессию ранних генов.

Для перехода фага в лизогенное состояние (и для последующего поддержания лизогенного состояния) необходимо установление репрессора (то есть, репрессор должен синтезироваться «с нуля» и закрепиться на операторах для последующей стабилизации своего количества). Для установления репрессора существует промотор P_{RE}, активируемый продуктом гена *cII*. Продукт еще одного гена, *cIII*, стабилизирует белок cII (препятствует его протеолитической деградации). Белок Cro выключает экспрессию *cII* и *cIII*, предотвращая наступление лизогении, а репрессор, со своей стороны, работает против хода литического цикла. Выбор фага между лизисом и лизогенией зависит от белка (Cro, или репрессора), который завладел операторами. Поэтому важнейшим фактором исхода инфекции является стабильность белка cII в инфицированной клетке.

Литература

14.4 Литический каскад контролируется двумя способами

Обзор

Greenblatt, J., Nodwell, J. R., and Mason, S. W. (1993). Transcriptional antitermination. *Nature* 364, 401–406.

14.6 Немедленноранние и задержанноранние гены фага λ необходимы для лизогении и для литического цикла

Обзор

Ptashne, M. (2004). *The Genetic Switch: Phage Lambda Revisited.* Cold Spring Harbor, NY: Cold Spring Harbor Press.

14.8 Лизогения поддерживается белком-репрессором

Исследование

Pirrotta, V., Chadwick, P., and Ptashne, M. (1970). Active form of two coliphage repressors. *Nature* 227, 41–44.

Ptashne, M. (1967). Isolation of the lambda phage repressor. *Proc. Natl. Acad. Sci. USA* 57, 306–313.

Ptashne, M. (1967). Specific binding of the lambda phage repressor to lambda DNA. *Nature* 214, 232–234.

14.9 Репрессор и его операторы формируют область иммунности

Обзор

Friedman, D. I. and Gottesman, M. (1982). *Lambda II.* Cambridge, MA: Cell Press.

14.10 ДНК-связывающая форма репрессора представляет собой димер

Исследование

Pabo, C. O. and Lewis, M. (1982). The operator-binding domain of lambda repressor: structure and DNA recognition. *Nature* 298, 443–447.

14.11 Мотив «спираль-поворот-спираль» необходим репрессору для связывания ДНК

Исследование

Sauer, R. T. et al. (1982). Homology among DNA-binding proteins suggests use of a conserved super-secondary structure. *Nature* 298, 447–451.

14.12 «Спираль узнавания» задает специфичность к ДНК

Исследование

Brennan, R. G. et al. (1990). Protein-DNA conformational changes in the crystal structure of a lambda Cro-operator complex. *Proc. Natl. Acad. Sci. USA* 87, 8165–8169.

Wharton, R. L., Brown, E. L., and Ptashne, M. (1984). Substituting an α-helix switches the sequence specific DNA interactions of a repressor. *Cell* 38, 361–369.

14.13 Димеры репрессора связываются с оператором кооперативно

Исследование

Bell, C. E., Frescura, P., Hochschild, A., and Lewis, M. (2000). Crystal structure of the lambda repressor C-terminal domain provides a model for cooperative operator binding. *Cell* 101, 801–811.

Johnson, A. D., Meyer, B. J., and Ptashne, M. (1979). Interactions between DNA-bound repressors govern regulation by the phage lambda repressor. *Proc. Natl. Acad. Sci. USA* 76, 5061–5065.

14.14 Находящийся на O_R2 репрессор взаимодействует с РНК-полимеразой на P_{RM}

Исследование

Hochschild, A., Irwin, N., and Ptashne, M. (1983). Repressor structure and the mechanism of positive control. *Cell* 32, 319–325.

Li, M., Moyle, H., and Susskind, M. M. (1994). Target of the transcriptional activation function of phage lambda cI protein. *Science* 263, 75–77.

14.16 Кооперативный характер связывания повышает чувствительность регуляции

Обзор

Ptashne, M. (2004). *The Genetic Switch: Phage Lambda Revisited.* Cold Spring Harbor, NY: Cold Spring Harbor Press.

Исследование

Bell, C. E. and Lewis, M. (2001). Crystal structure of the lambda repressor C-terminal domain octamer. *J. Mol. Biol.* 314, 1127–1136.

Dodd, I. B., Perkins, A. J., Tsemitsidis, D., and Egan, J. B. (2001). Octamerization of lambda CI repressor is needed for effective repression of P(RM) and efficient switching from lysogeny. *Genes Dev.* 15, 3013–3022.

Репликон

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

15.1 Введение

15.2 Репликоны бывают линейными и кольцевыми

- Реплицированный участок ДНК, фланкированный нереплицированной ДНК, похож на глаз.
- Репликативная вилка образуется в ориджине и передвигается вдоль ДНК.
- Если в ориджине образовалась одна репликативная вилка, репликация получается однонаправленной.
- Если в ориджине образовались две репликативные вилки, которые движутся в противоположных направлениях, репликация является двунаправленной.

15.3 Ориджины можно картировать при помощи авторадиографии и электрофореза

- Движение репликативной вилки можно детектировать при помощи авторадиографии после мечения в импульсном режиме.
- Репликативные вилки придают молекулам ДНК Y-образную форму, что меняет их электрофоретическую подвижность.

15.4 Влияет ли метилирование ориджина на инициацию репликации?

- Последовательность oriC содержит одиннадцать $\frac{GATC}{CTAG}$ повторов, метилированных по аденинам обеих цепей.
- Первичный продукт репликации — это наполовину метилированная ДНК; репликация такой ДНК не может быть инициирована.
- Приведение повторов $\frac{GATC}{CTAG}$ в полностью метилированное состояние (пригодное для инициации) выполняется после тринадцатиминутной задержки.

15.5 После репликации ориджины могут быть изолированы от аппарата обеспечения их функции

- Белок SeqA связывает наполовину метилированную ДНК; этот белок необходим, чтобы задержать повторную репликацию.
- SeqA может взаимодействовать с DnaA.
- В период, когда ориджины метилированы наполовину, они взаимодействуют с клеточной мембраной и могут быть недоступны для метилаз.
- Как происходит взаимодействие ориджина с мембраной, пока неясно.

15.6 Эукариотические хромосомы состоят из множества репликонов

- Один эукариотический репликон может иметь длину от 40 до 100 тпн.
- Каждая хромосома содержит множество репликонов.
- В течение S-фазы клеточного цикла каждый из этих репликонов активируется в индивидуальном порядке и в свое характерное время.
- Паттерны репликативной активности, снятые для обширных областей генома, свидетельствуют, что репликоны, близкие друг к другу, склонны активироваться в одно и то же время.

15.7 Выделение ориджинов из генов дрожжей

- У дрожжей *S. cerevisiae* ориджины представляют собой короткие AT-богатые участки ДНК, содержащие последовательность длиной 11 пн, существенную для их функции.
- ORC — это комплекс из шести белков, который связывается с *ARS*.

15.8 Лицензирующие факторы контролируют репликацию геномов эукариот

- Для инициации репликации любому ориджину необходим лицензирующий фактор.
- Эти факторы присутствуют в ядре до репликации; во время репликации они инактивируются или разрушаются.
- Инициация следующего раунда репликации возможна не раньше момента, когда ядро получит из цитоплазмы новую порцию лицензирующих факторов; а такой момент наступает только в митозе.

15.9 Лицензирующий фактор состоит из белков MCM

- У дрожжей белковый комплекс ORC ассоциирован с ориджинами репликации во всех фазах клеточного цикла.
- Нестабильный белок Cdc6 синтезируется только в фазе G1.
- Белок Cdc6, связываясь с ORC, разрешает белкам MCM также связаться с ORC.
- После того как репликация инициирована, белки Cdc6 и MCM уходят; деградация Cdc6 препятствует повторной инициации.
- Некоторые белки MCM присутствуют в ядре всегда, остальные могут попасть в ядро только во время митоза.

15.10 В митохондриях ориджины репликации имеют вид D-петель

- Для инициации репликации двух цепей мтДНК используются отдельные ориджины с разными последовательностями.
- Репликация H-цепи инициируется в D-петле.
- Репликация L-цепи инициируется, когда первая репликативная вилка достигнет ее ориджина.

15.11 Резюме

15.1 Введение

Вне зависимости от того, сколько хромосом содержит клетка, перед каждым делением ее геном должен быть полностью реплицирован, причем ровно один раз. Каким образом акт репликации связан с клеточным циклом?

Для соотнесения процесса репликации с ходом клеточного цикла клетка использует два общих принципа:

- *Инициация репликации неизбежно влечет за собой деление клетки.* Можно сказать, что число клеточных потомков определяется серией принятых клеткой решений о том, инициировать репликацию или нет. Репликация контролируется на стадии инициации. *Стоит репликации начаться, и она будет идти до тех пор, пока весь геном не окажется дуплицированным.*
- Если репликация началась, до ее завершения последующее деление не может произойти. Завершение репликации служит сигналом для клеточного деления. При делении дуплицированные геномы распределяются (сегрегируют) по одному в каждую дочернюю клетку. Единицей сегрегации служит хромосома.

У бактерий инициация репликации — это одно событие, в котором участвует уникальная область хромосомы. Деление бактериальной клетки происходит за счет образования септы — перегородки, растущей от клеточной стенки внутрь. В клеточном цикле эукариот инициация репликации (которая представляет собой множество отдельных событий) связана с началом S-фазы — продолжительного периода синтеза ДНК. При делении эукариотическая клетка основательно реорганизуется; эти изменения, включая само деление, входят в понятие «митоз».

В этой главе нас будут занимать вопросы, касающиеся регуляции репликации ДНК. Как инициируется раунд репликации? Каким образом отслеживается и контролируется его прогресс? Что говорит о его завершении?

Единицу ДНК, являющуюся объектом одного акта репликации, называют **репликоном**. В каждом клеточном цикле каждый репликон «выстреливает» один и только один раз. Репликон можно опознать по наличию в нем регуляторных элементов, необходимых для репликации. В нем есть **ориджин** — область, где инициируется репликация. В нем также может быть **участок терминации**, на котором репликация останавливается.

Любая последовательность, закрепленная за ориджином (в силу того что находится в той же молекуле ДНК, и между ней и ориджином нет участка терминации), будет реплицирована под контролем этого ориджина, как часть данного репликона. Ориджин взаимодействует со своим репликоном в *цис*-манере. Это значит, что он контролирует репликацию только той ДНК, частью которой является сам.

Исходно определение репликона для прокариот звучало как «область ДНК, в которой есть *не только* ориджин, *но и* ген, кодирующий белок-регулятор». В наши дни термин «репликон» в равной мере используют применительно к эукариотическим хромосомам для обозначения единицы репликации, которая содержит ориджин; белки-регуляторы, действующие в *транс*-манере, могут кодироваться в другом месте.

Геном бактериальной клетки укладывается в одну хромосому и представляет собой один-единственный репликон, так что единица репликации и единица сегрегации в данном случае совпадают. Событие инициации в уникальном ориджине равносильно запуску репликации целого генома; на каждое клеточное деление приходится ровно одно такое событие. Поскольку каждая гаплоидная бактерия имеет одну хромосому, репликация контролируется по типу **единственной копии** (однокопийно).

Помимо их собственной хромосомы, бактериальные клетки могут содержать дополнительную генетическую информацию в форме плазмид. Каждая плазмидная копия представляет собой автономную кольцевую ДНК, являющуюся репликоном. Этот репликон может, подобно бактериальной хромосоме, контролироваться по типу единственной копии (тогда он будет реплицироваться не чаще и не реже, чем хромосома бактерии-хозяина) или может контролироваться по типу **множества копий** (многокопийно). Во втором случае число плазмидных копий превышает число хромосом.

ДНК вирусов и бактериофагов тоже являются репликонами. В течение одного инфекционного цикла их репликация может инициироваться многократно. С учетом этого факта классическое определение репликона вырождается в утверждение, что *любая молекула ДНК, содержащая ориджин, может быть автономно реплицирована в клетке.*

Принципиальное различие между бактериальным и эукариотическим геномами просматривается в том, как эти геномы реплицируются. Каждая эукариотическая хромосома содержит большое число репликонов; то есть одна единица сегрегации содержит множество единиц репликации. Это обстоятельство придает проблеме контроля репликации генома еще большую глубину. С одной стороны, все репликоны хромосомы должны активироваться в течение клеточного цикла; с другой, они не делают этого одновременно. С одной стороны, на обязательную активацию всех репликонов отводится довольно-таки продолжительный интервал клеточного цикла; с другой стороны, повторная активация одних и тех же репликонов в одном клеточном цикле *совершенно недопустима.* Должен существовать какой-то сигнал, исключающий возможность повторной активации. Многие репликоны активируются независимо, поэтому должен существовать дополнительный сигнал — о том, что во всех репликонах процесс репликации полностью завершился.

К настоящему времени накоплено довольно много информации о том, как устроены отдельные эукариотические репликоны, а вот информации о том, как эти репликоны соотносятся, пока недостаточно. Не известно, насколько точно воспроизводятся паттерны репликации в каждом клеточном цикле. Все ли ориджины используются или некоторые из них иногда молчат? Всегда ли ориджины «выстреливают» в одном и том же порядке? Если существуют разные типы ориджинов, *что* отличает их друг от друга?

ДНК митохондрий и хлоропластов по типу регуляции похожи на плазмиды, существующие во множестве копий на одну бактерию — в отличие от хромосомной ДНК ядра, воспроизводство которой регулируется по типу единственной копии. ДНК хлоропластов и митохондрий представлена в клетке (и даже в органелле) множеством копий; воспроизводство этой ДНК должно быть тоже как-то соотнесено с клеточным циклом.

В отношении любой ДНК, если она является репликоном и интересует нас как репликон, перед исследователями стоят две задачи. Во-первых, определить последовательности, которые функционируют как ориджины; во-вторых, понять, почему и как их узнают специализированные белки аппарата репликации. Мы начнем эту главу с рассмотрения основ устройства репликонов и различных форм, которые они принимают; а также с биологической сущности ориджина. После этого мы перейдем к вопросу о том, как репликация бактериального генома оказывается скоординирована с клеточным циклом и от чего зависит сегрегация геномов по дочерним бактериальным клеткам.

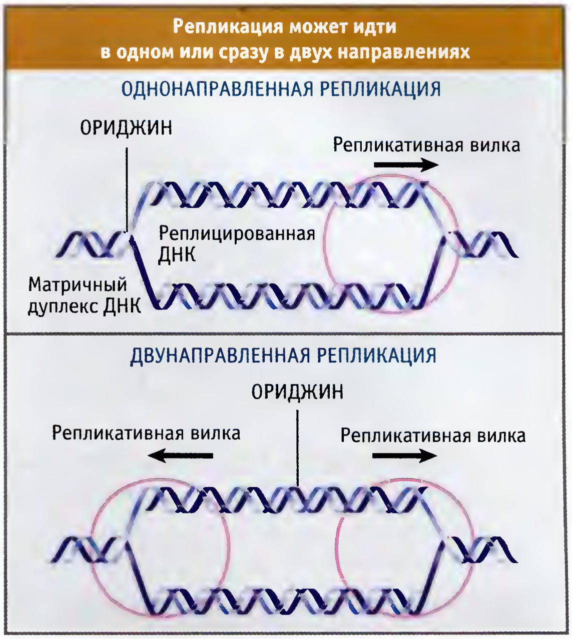


РИС. 15.2 Репликон может быть ориентирован в одном направлении или в обоих направлениях; от этого зависит, одна или две репликативные вилки формируются в ориджине на старте каждого раунда репликации

15.2 Репликоны бывают линейными и кольцевыми

Ключевые понятия:

- Реплицированный участок ДНК, фланкированный нереплицированной ДНК, похож на глаз.
- Репликативная вилка образуется в ориджине и передвигается вдоль ДНК.
- Если в ориджине образовалась одна репликативная вилка, репликация получается однонаправленной.
- Если в ориджине образовались две репликативные вилки, которые движутся в противоположных направлениях, репликация является двунаправленной.

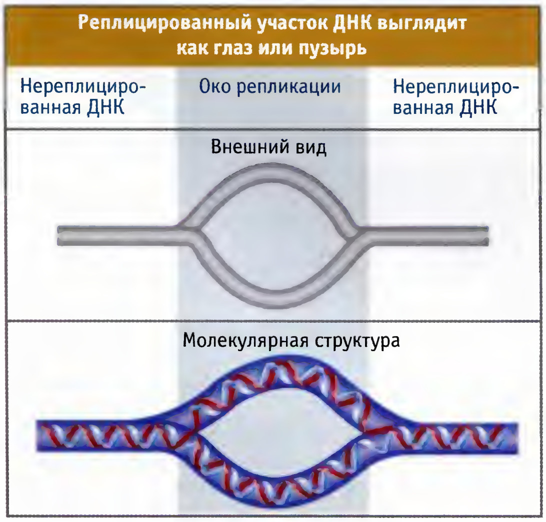


РИС. 15.1 Реплицированная ДНК образует око репликации, фланкированное еще не реплицированной ДНК

В реплицирующей молекуле ДНК есть два типа участков. Если посмотреть в электронный микроскоп, уже реплицированный участок будет похож на глаз (см. РИС. 15.1). Этот участок так и называют — **око репликации**. С обеих сторон око окружают еще не реплицированные участки ДНК, состоящие из матричного дуплекса. Око репликации состоит из раскрытого матричного дуплекса, на котором уже образовались два дочерних дуплекса.

Точка, в которой *идет* репликация, находится на границе участков и называется **репликативной вилкой**, или **точкой роста**. *Репликативная вилка начинает свое последовательное движение по ДНК от стартовой точки в ориджине.* Ориджин может быть использован для старта **однонаправленной репликации** или **двунаправленной репликации**. Если из ориджина вышла одна репликативная вилка, то репликация однонаправленная. Если из ориджина вышли, в противоположных направлениях, две репликативные вилки, то репликация является двунаправленной.

Независимо от того, одна там вилка или две, око репликации будет выглядеть одинаково. Если око образуется в ходе однонаправленной репликации, у него есть один фиксированный край (это ориджин) и один движущийся (это репликативная вилка). Если око образуется в ходе двунаправленной репликации, оба края (репликативные вилки) движутся, так что око» растет в обоих направлениях (РИС. 15.2). В любом случае репликация увеличивает око до тех пор, пока, в конце концов, оно не охватит весь репликон. Если репликон кольцевой, в ходе репликации он ока-



РИС. 15.3 Око репликации вместе с еще не реплицированным остатком кольцевой ДНК составляют θ-структуру (тэта-структуру)

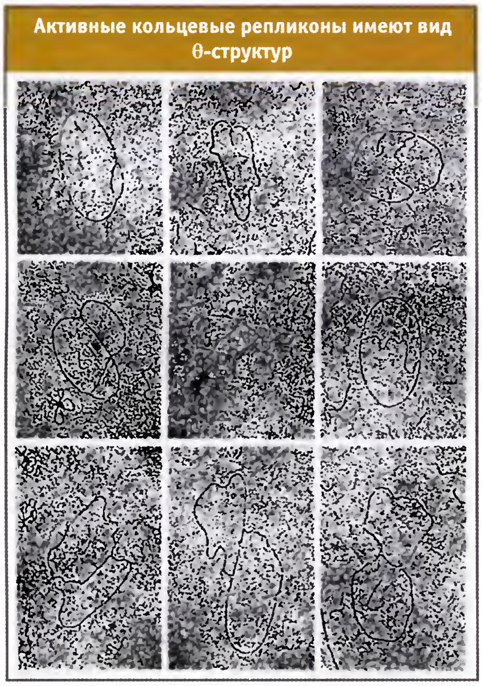


РИС. 15.4 По мере того как репликативные вилки обходят репликон, око репликации разрастается. Наступает момент, когда око становится больше, чем нереплицированный сегмент. Половинки ока имеют равную длину, поэтому их почти всегда можно отличить от нереплицированного сегмента. Фотография любезно предоставлена Bernard Hirt, Swiss Institute for Experimental Cancer Research (ISREC).

зывается похож на греческую букву тэта. Так называемая «θ-структура» изображена на РИС. 15.3. В конце репликации она становится похожа на очки. В качестве примера очень показательны последовательные стадии репликации кольцевой ДНК полиомавируса. Их электронные микрофотографии представлены на РИС. 15.4.

15.3 Ориджины можно картировать при помощи авторадиографии и электрофореза

Ключевые понятия:

- Движение репликативной вилки можно детектировать при помощи авторадиографии после мечения в импульсном режиме.
- Репликативные вилки придают молекулам ДНК Y-образную форму, что меняет их электрофоретическую подвижность.

Существует два способа определить, сколько точек роста. Выбор метода зависит от того, представлена ДНК отдельной молекулой или неким регионом клеточного генома.

Если мы имеем дело с определенной линейной молекулой, можно, используя электронную микроскопию, измерить расстояния от краев ока репликации до концов ДНК, а затем сравнить положения краев в молекулах с очами разных размеров. Если репликация однонаправленная, двигаться относительно концов ДНК будет только один край; второй край фиксирован (это ориджин). Если репликация двунаправленная, двигаться к концам будут обе вилки репликации; а ориджин будет расположен посередине.

Для исследования неопределенного фрагмента большого генома можно использовать радиоактивное мечение в импульсном режиме. Подача радиоактивного нуклеотида-предшественника в режиме дискретных импульсов позволяет детектировать смещение репликативных вилок за каждый временной интервал между подачей метки. Если делать импульсы разными по интенсивности, их следы можно идентифицировать по интенсивности включения метки в ДНК (то есть степени затемнения на радиоавтографе). На РИС. 15.5 показано, что при однонаправленной репликации череда затемнений появляется только у одного из краев ока. При дву-

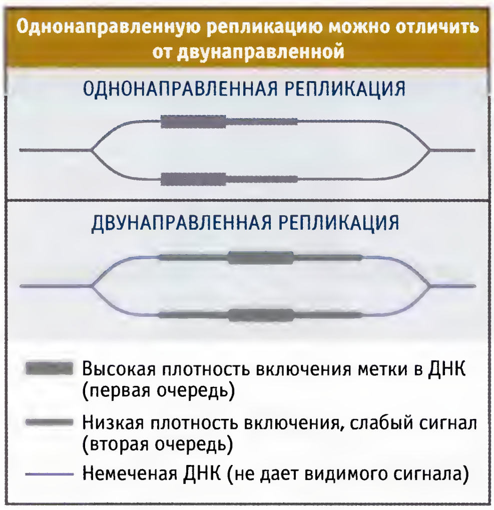


РИС. 15.5 Радиоактивное мечение, дифференцированное по интенсивности во времени, помогает отличить однонаправленную репликацию от двунаправленной

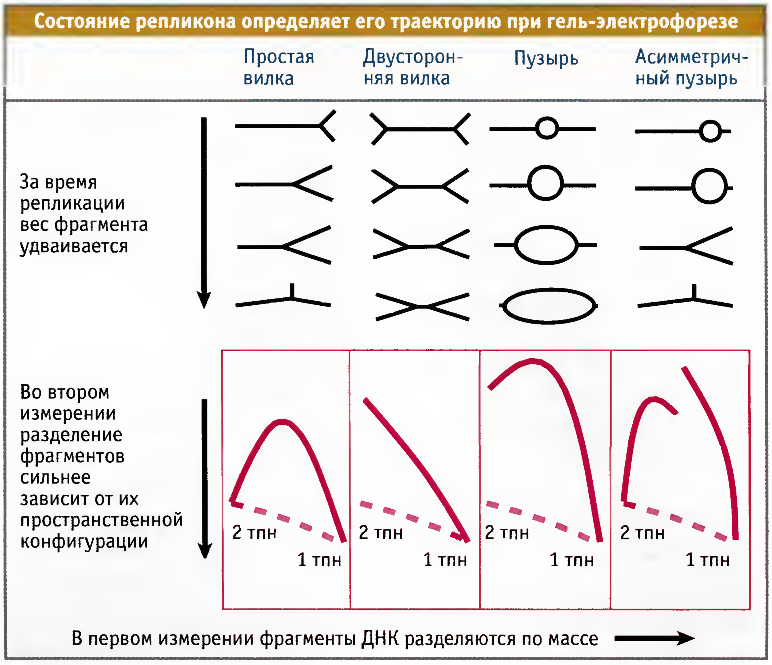


РИС. 15.6 Положение ориджина и число репликативных вилок задают структуру реплицирующегося фрагмента ДНК, которую можно отследить по геометрическому месту на электрофореграмме (сплошная линия). Пунктирной линией показано геометрическое место фрагментов линейной ДНК, не имеющих выраженной конфигурации

направленной репликации на «оке» возникает симметричный паттерн мечения, свойственный, в частности, репликонам эукариотических хромосом.

Более современный метод картирования ориджинов, отличающийся лишь более высоким разрешением, основан на том, что изменение геометрии молекулы ДНК влияет на ее электрофоретическую подвижность. На РИС. 15.6 изображена техника двумерного гель-электрофореза. Фрагменты ДНК (продукты расщепления эндонуклеазой рестрикции) «разгоняют» между электродами в одном измерении, где они разделяются по массе, и в другом измерении, где их движение в большей степени определяется формой. Разные виды реплицирующихся молекул образуют характерные паттерны своего местонахождения в форме кривых. Отклонение этих кривых от прямой линии, которая характеризует линейные молекулы (размер которых плавно возрастает от исходного до вдвое большего, поскольку при репликации масса ДНК удваивается), можно точно измерить и истолковать.

Простая Y-образная структура, в которой одна репликативная вилка движется вдоль линейного фрагмента, даст непрерывную кривую. Точка перегиба кривой соответствует моменту (точнее, молекулам, захваченным в тот момент), когда длины всех трех ветвей (двух дочерних и одного матричного дуплексов) равны между собой — из-за этого при электрофорезе вся структура сильнее всего отклоняется от прямой линии. По аналогии, можно интерпретировать кривые для двойных Y-образных структур и для так называемых «пузырей». Асимметричные пузыри образуют на электрофорезе разорванную кривую. Разрыв соответствует моменту, когда пузырь превращается в Y-образную структуру из-за того, что одна репликативная вилка добегает до конца матричного дуплекса ДНК.

Применение разных методик для описания реплицирующейся ДНК показало, что чаще всего ориджины используются для инициации *двунаправленной* репликации. Теперь мы попробуем взглянуть на репликацию при еще более высоком разрешении. В следующих разделах нам предстоит знакомство с регуляторными элементами, из которых состоит ориджин, и с *транс*-действующими факторами, которые узнают этот ориджин.

15.4 Влияет ли метилирование ориджина на инициацию репликации?

Ключевые понятия:

- Последовательность *oriC* содержит одиннадцать повторов $\frac{\mathrm{GATC}}{\mathrm{CTAG}}$, метилированных по аденинам обеих цепей.
- Первичный продукт репликации — это наполовину метилированная ДНК; репликация такой ДНК не может быть инициирована.
- Приведение повторов $\frac{\mathrm{GATC}}{\mathrm{CTAG}}$ в полностью метилированное состояние (пригодное для инициации) выполняется после тринадцатиминутной задержки.

Где гарантия, что ориджин бактериальной хромосомы (или плазмидной ДНК) будет использован для репликации ровно *один* раз на каждый клеточный цикл? Может быть, инициация репликации сопряжена с изменением в структуре ориджина, и это изменение маркирует ориджин как уже реплицированный?

Для этих целей используются некоторые последовательности, включаемые в ориджин. У *E. coli* в ориджине *oriC* содержатся одиннадцать копий последовательности $\begin{matrix}\text{GATC}\\\text{CTAG}\end{matrix}$. Это палиндромы GATC; в последовательности *oriC* их насчитывается одиннадцать. Каждый сайт GATC служит мишенью для фермента, который вносит в ДНК химическую модификацию. Этот элемент — аденин(N^6)-ДНК-метилтрансфераза(Dam) — метилирует аденин по атому азота в позиции N^6. Реакция изображена на РИС. 15.7.

До репликации сайт GATC метилирован по аденинам обеих цепей ДНК. Во время репликации новые цепи ДНК строятся из нуклеотидов с обычными (немодифицированными) азотистыми основаниями. Поэтому новосинтезированная цепь дочернего дуплекса *неметилирована*, так что дочерние дуплексы ДНК метилированы лишь наполовину — по матричной цепи. Можно сказать, что событие репликации переводит сайты-мишени Dam из полностью метилированного состояния в полуметилированное состояние.

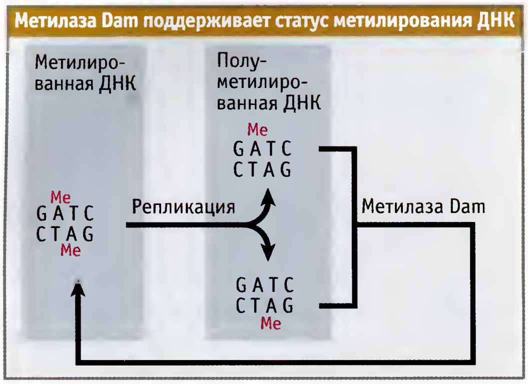


РИС. 15.7 Репликация метилированной ДНК дает наполовину метилированную ДНК; этот статус сохраняется за ДНК до тех пор, пока метилаза Dam не восстановит ее исходное, полностью метилированное состояние

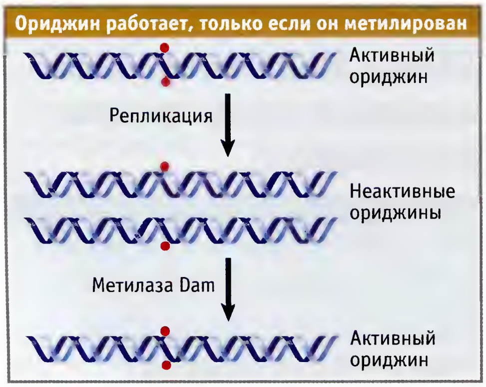


РИС. 15.8 Только в полностью метилированном ориджине может быть инициирована репликация. Наполовину метилированные дочерние ориджины не используются до тех пор, пока не будет восстановлено их полностью метилированное состояние

Оказывается, это обстоятельство имеет свой регуляторный смысл. Способность произвольной ДНК (например, плазмиды) с *oriC* реплицироваться в *dam⁻ E. coli* зависит от статуса метилирования. Если она метилирована (по аденинам GATC), то проходит один-единственный раунд репликации. Продукты этого первого раунда репликации так и останутся полуметилированы (потому что метилаза Dam в мутантах *dam⁻* не работает, и *добавочного* метилирования не происходит). Как выяснилось, репликация не может быть инициирована в ориджине, который метилирован наполовину (см. РИС. 15.8).

Можно предположить, что либо полностью метилированное состояние сайтов-мишеней Dam в *oriC* стимулирует инициацию, либо наполовину метилированное состояние этих сайтов ингибирует инициацию. Истинно, по-видимому, второе утверждение: оказалось, что *совершенно не метилированный* ориджин *oriC* вполне эффективен.

Итак, в наполовину метилированном ориджине *E. coli* репликация не может быть инициирована — до тех пор, пока метилаза Dam не превратит этот ориджин в полностью метилированный. Это произойдет через ~13 минут после репликации, что неестественно долго: в других частях генома *E. coli* добавочное метилирование сайтов GATC начинается почти сразу (меньше, чем через полторы минуты) после репликации. Между прочим, кроме ориджина в геноме *E. coli* есть еще одна область, где добавочное метилирование происходит с задержкой — это промотор гена *dnaA*.

Наполовину метилированный промотор *dnaA* не активен, и пока не произойдет добавочного метилирования, ген будет молчать; количество белка DnaA, ключевого инициатора репликации, при этом уменьшится. Таким образом, продукция DnaA тормозится одновременно с задержкой добавочного метилирования ориджина — в период, когда ориджин инертен.

15.5 После репликации ориджины могут быть изолированы от аппарата обеспечения их функции

Ключевые понятия:

- Белок SeqA связывает наполовину метилированную ДНК; этот белок необходим, чтобы задержать повторную репликацию.
- SeqA может взаимодействовать с DnaA.
- В период, когда ориджины метилированы наполовину, они взаимодействуют с клеточной мембраной и могут быть недоступны для метилаз.
- Как происходит взаимодействие ориджина с мембраной, пока неясно.

Чем же обусловлена задержка добавочного метилирования ДНК в *oriC* и *dnaA*? Вероятнее всего, в период задержки эти области хромосомной ДНК *E. coli* отстранены от своей функции («секвестированы») и находятся в форме, недоступной действию метилазы Dam.

Анализ фенотипов мутантов по одному из генов *E. coli*, *seqA*, помог создать схему, по которой идет контроль повторного использования ориджинов. У мутантов по *seqA* сокращена задержка добавочного метилирования *oriC* и *dnaA*; из-за этого повторная репликация ДНК инициируется слишком скоро, и хромосомные копии (неполные) накапливаются в избытке. Такой фенотип свидетельствует, что *seqA* входит в систему, препятствующую добавочному метилированию ориджинов.

Как выяснилось, белок SeqA крепче связывается с наполовину метилированной ДНК, чем с ДНК, которая метилирована полностью. И если связывание SeqA с ДНК инициируется в момент, когда ДНК становится наполовину метилированной (то есть сразу после репликации), то его непрерывное присутствие в ориджине будет препятствовать образованию открытого комплекса инициации. У белка SeqA нет специфичности к последовательности *oriC*; возможно, SeqA обретает специфичность к *oriC*, взаимодействуя с белком DnaA. Это хорошо объясняло бы генетические взаимодействия между *seqA* и *dnaA* (случаи супрессии между мутациями в этих генах).

Для ассоциации ориджина с клеточной мембраной *in vitro* необходимо, чтобы последовательности GATC в этом ориджине были наполовину метилированы. Наполовину метилированная ДНК *oriC* способна связываться с мембранами, а полностью метилированная — нет. Можно предположить, что ассоциация ориджина с мембраной вовлечена в регуляцию его активности (независимо от роли, которую играет мембрана в сегрегации бактериальных хромосом; см. рис. 17.4). Ассоциация с мембраной может препятствовать преждевременной повторной инициации либо напрямую ингибируя реакцию инициации какими-то компонентами мембраны, либо косвенно секвестируя ориджин от реакции.

Свойства мембранной фракции клеточных гомогенатов *E. coli* говорят о присутствии компонентов, регулирующих репликацию. Ингибитор, обнаруженный в этой фракции, конкурирует с белком DnaA. Этот ингибитор может предотвратить инициацию репликации *in vitro*, если он добавлен в систему прежде белка DnaA. На этом факте строится схема, изображенная на РИС. 15.9: ингибитор узнает наполовину метилированную ДНК и мешает DnaA с ней связаться. Добавочное метилирование ДНК провоцирует уход ингибитора. Если ингибитор находится в мембране, ассоциацию и диссоциацию ДНК с мембраной можно рассматривать как часть контроля репликации.

Полное представление о системе контроля повторной инициации репликации бактериальной хромосомы пока отсутствует. В ней может быть задействовано сразу несколько механизмов, обеспечивающих отстранение ориджина, задержку повторного метилирования, подавление взаимодействия DnaA с ДНК, репрессию транскрипции DnaA. Какой из эффектов первичен по отношению к остальным и насколько прямо или косвенно они влияют на инициацию? Какой конкретно механизм, на самом деле, служит «часами» для инициации? Возможно, при инициации репликации происходит прикрепление ориджина к мембране, и для его последующего освобождения требуется сборка какой-то крупной структуры? Длительность прикрепления ориджина к мембране возрастает с увеличением продолжительности клеточного цикла и, следовательно, имеет прямое отношение к скорости возникновения повторной инициации.

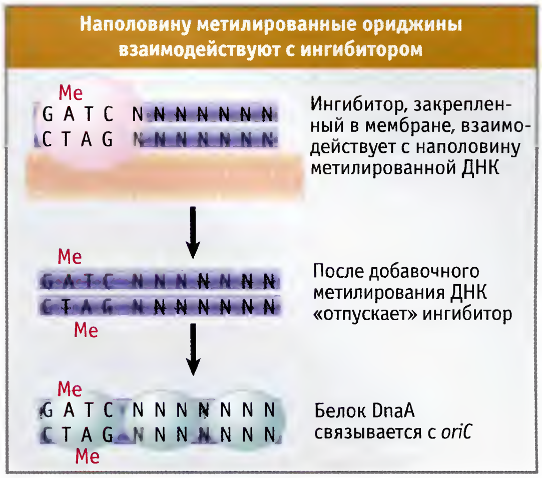


РИС. 15.9 Ингибитор, закрепленный в клеточной мембране, взаимодействует с наполовину метилированной ДНК ориджина; из-за него DnaA не может связаться с ориджином. Добавочное метилирование ДНК выводит ее из-под влияния ингибитора

Белок DnaA — компонент аппарата репликации, и, в отличие от остальных компонентов, востребованный только в ориджине. Этот белок служит мишенью для нескольких регуляторных систем. Возможно, ни одна из этих систем не способна самостоятельно контролировать частоту событий инициации. Некоторые мутации в *dnaA* вызывают разобщение «графика» инициаций с ходом клеточного цикла, поэтому DnaA можно представить себе в роли «титратора», измеряющего число ориджинов по отношению к клеточной массе. В то же время, перепроизводство клеткой этого белка дает крайне противоречивые результаты: от полного отсутствия какого-либо эффекта до преждевременной инициации.

Идентификация белков, опосредующих прикрепление ориджина к мембране, оказалась трудной задачей. Возможно, прикрепление ориджина к мембране — это одна из функций DnaA. Именно так можно интерпретировать характер взаимодействия DnaA с фосфолипидами *in vitro*. Если с DnaA связан АДФ, фосфолипиды стимулируют его обмен на АТФ. Не понятно, какую роль это играет в контроле активности DnaA, однако сам факт говорит о склонности DnaA взаимодействовать с мембраной. Можно предположить, что ассоциация ДНК с клеточной мембраной — это результат комбинации нескольких молекулярных событий. Наполовину метилированный

ориджин связывается с ингибитором, сидящим на мембране, а когда ориджин становится метилирован полностью, ингибитор вытесняется из него белком DnaA, также ассоциированным с мембраной.

Если белок DnaA действительно является инициатором, запускающим раунд репликации, то принципиальным этапом инициации должно быть его накопление в области ориджина до некоторого критического уровня. В общей внутриклеточной концентрации DnaA и в экспрессии *dnaA* не наблюдается значительных циклических вариаций. Тем не менее это свидетельствует в пользу того, что накопление DnaA в ориджине является локальным событием. Кроме того, чтобы инициировать репликацию, DnaA должен находиться в связанной с АТФ форме. У DnaA есть собственная слабая активность перевода АТФ в АДФ. Эта активность усиливается субъединицей β ДНК-полимеразы III. Когда репликаза (холофермент ДНК-полимеразы III) включается в состав комплекса репликации, взаимодействие DnaA с субъединицей β вызывает гидролиз АТФ, связанного с DnaA. Инактивированный DnaA не может инициировать следующий раунд репликации. Эта реакция известна как RIDA (от *англ.* regulatory inactivation of DnaA — регуляторная инактивация DnaA). Реакцию усиливает еще один белок, Hda. Чем определяется время, через которое наступает повторная активация DnaA, пока неясно.

Еще один фактор контроля доступа DnaA к связыванию ориджина — конкуренция со стороны других областей бактериальной хромосомы. Так, в локусе *dat* сосредоточено множество сайтов связывания DnaA; все вместе они могут связать в восемь раз больше DnaA, чем ориджин. Делеция *dat* повышает частоту событий инициации в ориджине. Наличие в хромосоме локуса *dat* значительно снижает количество DnaA, приходящееся на ориджин, хотя точный смысл этого снижения для своевременной инициации репликации пока неясен.

15.6 Эукариотические хромосомы состоят из множества репликонов

Ключевые понятия:

- Один эукариотический репликон может иметь длину от 40 до 100 тпн.
- Каждая хромосома содержит множество репликонов.
- В течение S-фазы клеточного цикла каждый из этих репликонов активируется в индивидуальном порядке и в свое характерное время.
- Паттерны репликативной активности, снятые для обширных областей генома, свидетельствуют, что репликоны, близкие друг к другу, склонны активироваться в одно и то же время.

В эукариотической клетке репликация ядерной ДНК приурочена к определенному интервалу клеточного цикла. В клетках высших эукариот этот интервал обычно длится несколько часов. Его называют S-фазой. В S-фазе происходит репликация огромного количества ДНК, содержащегося в каждой из хромосом. Эта задача может быть выполнена за счет разбиения хромосомы на *множество* репликонов. В каждый конкретный момент S-фазы активны (то есть активно дуплицируются) не все, а только некоторые репликоны. Можно сказать, что у каждого репликона есть свое характерное время активации (с оговоркой, что не все имеющиеся данные убедительно об этом свидетельствуют).

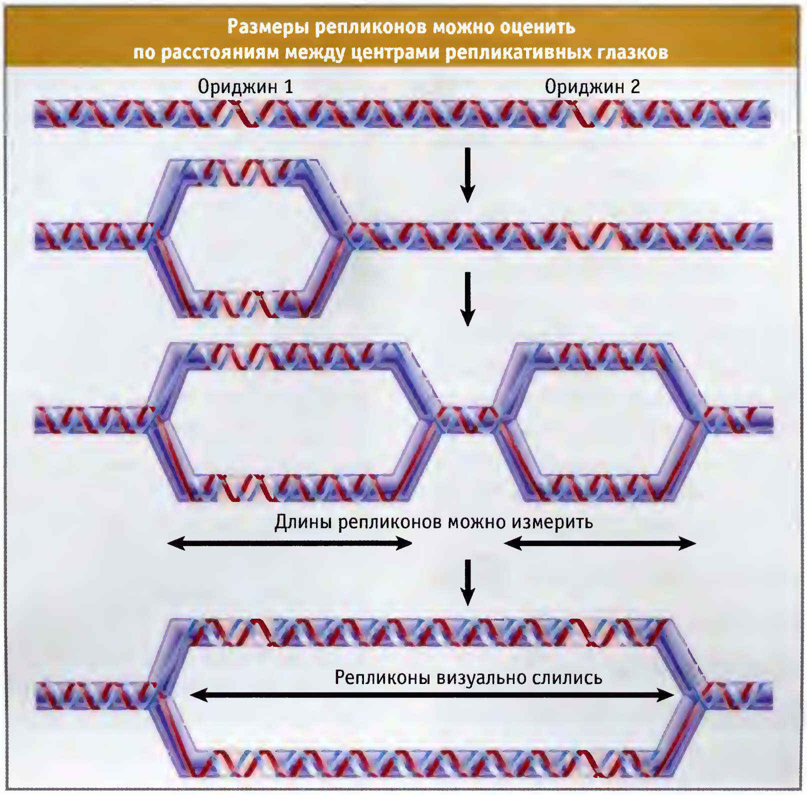
Начало S-фазы — это активация самых первых репликонов. В последующие несколько часов события инициации по порядку охватывают все остальные репликоны. Значительная часть того, что нам известно о свойствах эукариотических репликонов, получена в экспериментах с применением метода авторадиографии. Под «авторадиографией» подразумевается последовательность манипуляций, изображенных на рис. 15.5 и 15.6. Эксперименты показали, что, как правило, в репликонах эукариот имеет место двунаправленная репликация.

Насколько велик среднестатистический репликон? Сколько всего репликонов в одном геноме? Вычленить репликон как обособленную единицу довольно трудно. Проблема в том, что смежные репликоны сливаются друг с другом, образуя общее око; это изображено на РИС. 15.10. Чтобы избежать ошибок в интерпретации продуктов слияния (fused eyes), подвергать анализу следует только те участки ДНК, в которых удалось поймать активными сразу несколько репликонов, по возможности в момент после синхронной инициации репликации и до того, как репликативные вилки смежных репликонов успели встретиться.

В таких группах активных репликонов можно измерить средний размер репликона (как усредненное расстояние между центрами очей). Скорость движения репликативной вилки можно измерить как максимальную протяженность сигнала о включении в ДНК радиоактивного изотопа после сеанса мечения фиксированной продолжительности.

Как удалось выяснить, в геномах эукариот отдельные репликоны относительно невелики. У дрожжей и мух средний размер репликона составляет ~40 тпн, у млекопитающих ~100 тпн. (При этом внутри одного генома длина репликонов может различаться в десятки раз!) Скорость движения репликативной вилки составляет ~2000 пн в минуту (у бактерий скорость движения репликативной вилки гораздо выше, 50 000 пн в минуту). Зная скорость репликации и средний размер репликона, можно вычислить, что геном млекопитающего может быть реплицирован за ~1 час при условии, что все репликоны «выстрелят» одновременно.

Но на самом деле в типичной соматической клетке млекопитающего S-фаза длится более 6 часов. Если предположить, что процесс удвоения количества ДНК протекает равномерно, можно рассчитать, что в каждый момент времени активно не более 15% всех репликонов. Такая неспешность не всегда позволительна. Например, в раннем эмбриональном развитии *Drosophila* проме-



На участке, где активны смежные репликоны, можно измерить их длины

жутки между клеточными делениями очень короткие. Соответственно, S-фаза тоже предельно коротка. Это достигается за счет одновременной активности огромного числа репликонов.

Но все же в большинстве случаев репликоны «выстреливают» не одновременно. Каким образом происходит распределение огромного числа событий инициации на интервале времени S-фазы? Почему ориджины, или группы ориджинов, соблюдают при этом некоторую очередность?

У дрожжей *Saccharomyces cerevisiae* большинство ориджинов, по умолчанию, активируется в ранней S-фазе. Однако некоторые ориджины сообщаются с определенными *цис*-действующими последовательностями-тормозами. Для таких ориджинов старт может быть отложен на более поздний срок, ближе к концу S-фазы.

Все данные говорят о том, что в хромосомах эукариот репликоны не имеют специализированных участков терминации, в которых репликативные вилки прекращают свое движение и диссоциируют от ДНК. Скорее всего, репликативная вилка, вышедшая из ориджина, будет продолжать свое движение до тех пор, пока не столкнется с вилкой смежного репликона, движущейся во встречном направлении. Напомним также о потенциальной топологической проблеме соединения новосинтезированных ДНК на стыке репликонов (то есть на месте встречи репликативных вилок).

Рассмотрим два репликона. Чем ближе они находятся, тем выше вероятность, что они реплицируются одновременно. Эта тенденция соседних участков генома быть активными в одно и то же время может объясняться существованием «региональных» систем контроля, благодаря которым они «выстреливают» более или менее координированно. (В противоположность диффузной активации отдельных репликонов одного за



РИС. 15.11 В S-фазном ядре эукариотической клетки репликативные вилки образуют скопления, или фокусы. Клетки, представленные на фотографиях, были помечены BrdU. Фото слева демонстрирует окраску йодистым пропидием, выявляющую весь массив ДНК. На фото справа клетки были окрашены антителами к BrdU для выявления реплицирующейся ДНК. Фотография любезно предоставлена Anthony D. Mills and Ron Laskey, Hutchinson/MRC Research Center, University of Cambridge

другим в разрозненных областях генома.) О возможно большем масштабе организации контроля синхронной активации соседствующих репликонов говорят как минимум две структурных особенности эукариотической хромосомы. Во-первых, целые, очень большие, области хромосом можно охарактеризовать как рано реплицирующиеся или поздно реплицирующиеся. Получается, что репликоны, активные на ранних и поздних сроках S-фазы, между собой почти не перемешаны. Во-вторых, множество визуализированных мечением репликативных вилок не разбросано по хромосомам, а сконцентрировано в виде отдельных очагов или скоплений («foci»). Таких скоплений может быть от 100 до 300. Несколько скоплений показано на РИС. 15.11, каждый из них содержит более 300 репликативных вилок. Скоплениям могут соответствовать какие-нибудь структуры, сквозь которые должна двигаться реплицирующаяся ДНК.

15.7 Выделение ориджинов из геномов дрожжей

Ключевые понятия:

- У дрожжей *S. cerevisiae* ориджины представляют собой короткие АТ-богатые участки ДНК, содержащие последовательность длиной 11 пн, существенную для их функции.
- ORC — это комплекс из шести белков, который связывается с *ARS*.

Известно, что в эукариотических клетках плазмиды отсутствуют. Однако любой сегмент ДНК, имеющий точку начала репликации (ориджин), способен реплицироваться. Из этого следует, что путем соответствующих манипуляций *in vitro* можно сконструировать плазмиды, способные к существованию в эукариотических клеточных линиях. Эта задача была успешно решена для дрожжей, но не для высших эукариот.

Клетки *S. cerevisiae*, мутантные по какой-либо функции, могут быть трансформированы в клетки с диким фенотипом при помощи добавления ДНК, несущей копию гена дикого типа. Оказалось, что некоторые фрагменты геномной ДНК дрожжей, если они замкнуты в кольцо, способны трансформировать дефектные клетки очень эффективно. Такие фрагменты, продолжая оставаться *вне* хромосом, могут закрепиться в клеточной линии, то есть обрести статус автономно реплицирующихся плазмид. Их анализ позволил идентифицировать ориджины в геноме дрожжей.

Фрагмент ДНК, с высокой вероятностью трансформирующий все потомство мутантной клетки (high frequency transforming fragment), содержит особую последовательность, позволяющую эффективно реплицироваться в дрожжевой клетке. Эта последовательность получила название **ARS** (от *англ.* autonomously replicating sequence — автономно реплицирующаяся последовательность). Элементы *ARS* эквивалентны ориджинам репликации.

Картирование элементов *ARS* было выполнено на обширных областях генома дрожжей. Оказалось, что только некоторые элементы *ARS* на самом деле используются для инициации репликации; остальные молчат или, возможно, используются, но не в каждом клеточном цикле. Последнее утверждение означает, что соответствующие участки хромосомы будут в разных клеточных циклах реплицироваться с разных ориджинов. Отсюда неизбежно следует, что участки терминации между соответствующими репликонами не могут быть фиксированными.

В А-Т-богатом элементе *ARS* есть дискретные сайты, мутации по которым нарушают функцию ориджина. За исключением этих сайтов, в элементе *ARS* более важен состав оснований, чем их точная последовательность. На РИС. 15.12 приведены результаты последовательного анализа мутаций по всей длине элемента *ARS*. Полная утрата функции ориджина наступает в случаях мутаций в сердцевине ориджина — сайте в 14 пн, который называют **доменом А**. Сайт А содержит консенсус из 11 пар оснований А-Т. Сходство нуклеотидных последовательностей элементов *ARS* ограничивается этим консенсусом. Иногда его называют *ACS* (*ARS* consensus sequence).

Элементы *ARS* содержат еще три сайта, мутации по которым снижают эффективность ориджина. Их обозначаются как В1, В2 и В3. Ориджин может эффективно функционировать с любыми двумя из трех сайтов В при условии, что в нем присутствует сайт А. Рядом с каждым сайтом В, или перекрываясь с ним, расположены несовершенные копии консенсуса *ACS* (в позиции 9/11), но для функции ориджина это не принципиально.

Комплекс, который связывает сайты А и В1, состоит из шести белков общей массой 400 кДа и получил название ORC (от *англ.* origin recognition complex — комплекс узнавания ориджина). Комплекс ORC ассоциирован с *ARS* во всех фазах клеточного цикла, что означает, что инициация репликации может зависеть от изменений состояния ORC, но не от их с ориджином взаимодействия *de novo* (см. разд. 15.9 «Лицензирующий фактор состоит из белков МСМ»). Подсчет геномных сайтов, с которыми связан ORC, говорит о наличии в дрожжевой хромосоме около 400 ориджинов репликации. Это соответствует средней длине репликона ~35 тпн.

Копии ORC были найдены и у высших эукариот. Комплекс ORC впервые был обнаружен у дрожжей *S. cerevisiae* и получил название scORC — после того, как похожие комплексы были описаны у *Schizosaccharomyces pombe* (spORC), *Drosophila melanogaster* (DmORC) и *Xenopus laevis* (XlORC). Все они связывают ДНК, часто их локализация ассоциирована с инициацией репликации (хотя для остальных организмов сайты связывания не выявляются с такой точностью, как для *S. cerevisiae*). Таким образом, ORC — это комплекс инициации, связывание которого помогает идентифицировать точку начала (ориджин). Однако подробности взаимодействия ORC с ДНК известны только для *S. cerevisiae*. Вполне возможно, что у других организмов узнавание ДНК комплексом ORC требует участия дополнительных компонентов.

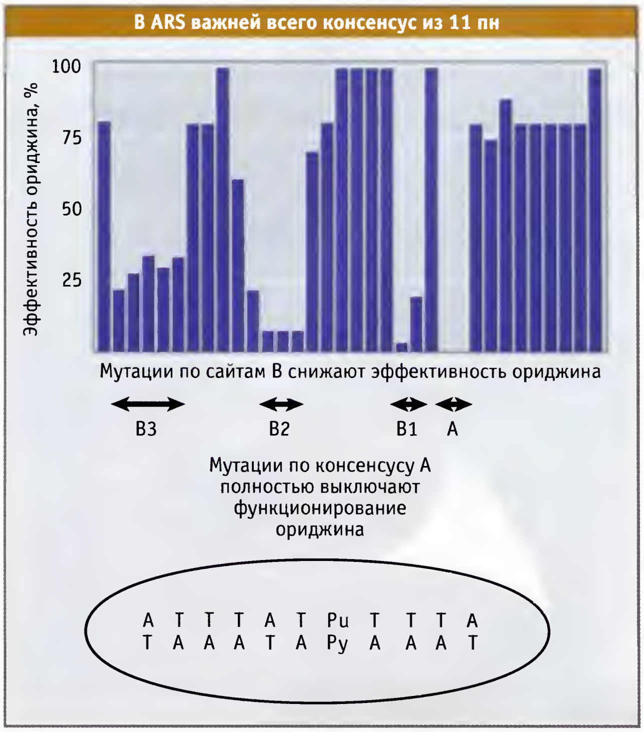


РИС. 15.12 Элемент ARS имеет длину ~50 пн; он содержит консенсус A и дополнительные сайты B1-B3

Элементы *ARS* у дрожжей *S. cerevisiae* удовлетворяют классическому определению ориджина (точки начала репликации) как последовательности, *цис*-действующей в качестве инициатора репликации ДНК. Последовательности элементов *ARS* мало похожи на последовательности предполагаемых ориджинов у других организмов. И все же консервативность комплексов ORC говорит о том, что ориджины высших эукариот работают по тем же принципам, что и ориджины дрожжей.

В ориджинах высших эукариот редко удается вычленить консенсус. Может быть, у высших эукариот ориджины устроены сложнее, чем у дрожжей? И к высшим эукариотам неприменима концепция ориджина как области, где есть дискретные *цис*-действующие последовательности, через которые идет полный контроль инициации репликации? Существует предположение, что в клетках животных некоторые репликоны имеют комплексные паттерны инициации: в некоторых случаях в водном регионе обнаруживается сразу несколько репликационных глазков. Соответственно, возникает вопрос об альтернативности и множественности ориджина и о том, существует ли ориджин как небольшая дискретная область генома.

Противоречие между консервативностью (то есть очевидной универсальностью) комплексов ORC и неопределенностью со способами их действия у разных организмов ослабло после обнаружения влияния условий среды на использование ориджинов. Для конкретной области генома, в которой характерно появление множества репликационных пузырьков синхронно инициированной репликации, может быть выявлен первичный ориджин. В условиях изобилия нуклеотидов-предшественников используется только он один. Если нуклеотидов-предшественников не хватает, то, помимо первичного ориджина, используется множество вторичных ориджинов (паттерн со множеством пузырьков характерен как раз для этого случая). По-видимому, в условиях нехватки нуклеотидов повторная инициация не может состояться быстро, комплекс ORC сходит с первичного ориджина и инициирует репликацию в другом месте, неподалеку. Скорее всего, последовательности-детерминанты ориджинов репликации у высших эукариот все-таки существуют, и в ближайшее время эти последовательности будут идентифицированы.

15.8 Лицензирующие факторы контролируют репликацию геномов эукариот

Ключевые понятия:

- **Для инициации репликации любому ориджину необходим лицензирующий фактор.**
- **Эти факторы присутствуют в ядре до репликации; во время репликации они инактивируются или разрушаются.**
- **Инициация следующего раунда репликации становится возможна не раньше момента, когда ядро получит из цитоплазмы новую порцию лицензирующих факторов; а такой момент наступает только в митозе.**

Эукариотический геном состоит из множества репликонов. В каждом клеточном цикле (интервале существования клетки от деления до деления) ориджин каждого репликона активируется один и только один раз. Каким же образом система добивается выполнения этого условия? Какие белки в этом участвуют? Выполнение условия может обеспечиваться присутствием в ориджине какого-нибудь фактора: одноразового активатора либо репрессора, запрещающего повторную активацию.

Замечательную модельную систему для слежения за ДНК, проходящей раунд репликации, представляет собой яйцеклетка шпорцевой лягушки *Xenopus laevis*. Яйцеклетка содержит все компоненты, необходимые для репликации: ведь первые несколько часов после оплодотворения зигота дробится, несмотря на то что ее гены молчат на протяжении первых одиннадцати клеточных циклов. Значит, если ядро какой-нибудь другой клетки инъецировать в яйцеклетку *X. laevis*, его ДНК может быть реплицирована. Некоторое представление о технике и логике соответствующих экспериментов дает РИС. 15.13.

Если в яйцеклетку посредством инъекции пересадить ядро сперматозоида (либо интерфазное ядро соматической клетки), ДНК этого ядра будет реплицирована только один раз. За этим легко проследить, используя мечение ДНК тяжелыми предшественниками (density label), которое изначально использовали для

доказательства полуконсервативной природы репликации (см. рис. 1.12). Если в такой яйцеклетке искусственно заблокировать синтез белка, то оболочка пересаженного ядра будет сохранять свою целостность, и его ДНК больше не будет реплицироваться. Если же синтез белка идет, в оболочке пересаженного ядра появляются разрывы, как будто бы клетка собралась делиться, и тогда в материале пересаженного ядра могут иметь место последующие раунды репликации ДНК. Того же эффекта можно добиться и в отсутствие синтеза белка — при помощи специальных химических агентов, делающих оболочку ядра проницаемой. Эти результаты означают, что в ядре есть какой-то компонент, нужный для репликации и полностью расходуемый за один раунд репликации. В цитоплазме яйцеклетки этот компонент присутствует постоянно, но, пока у ядра цела оболочка, она закрывает компоненту доступ в ядро.

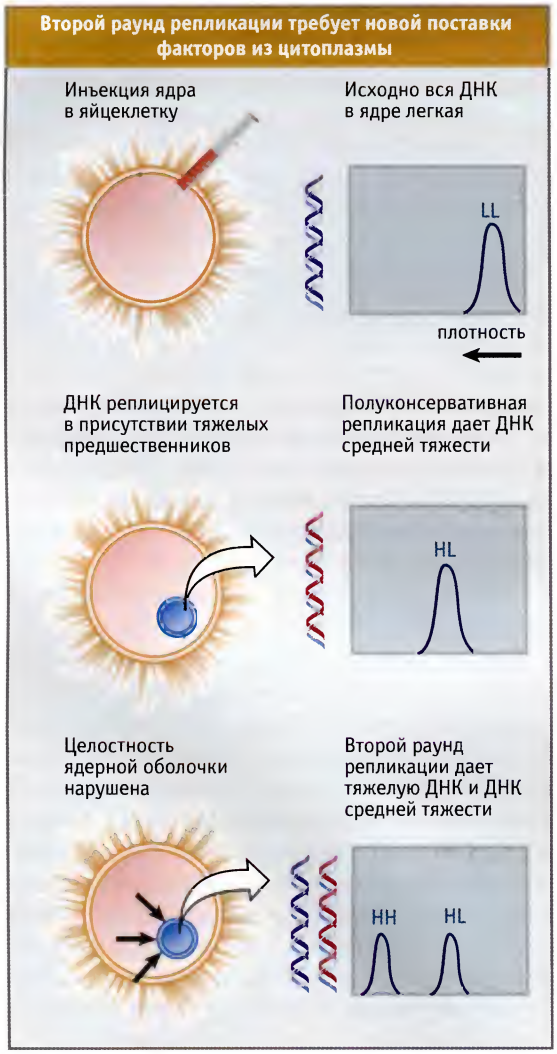


РИС. 15.13 Ядро, инъецированное в яйцеклетку *Xenopus*, может реплицировать свою ДНК только один раз. Для того чтобы состоялись последующие раунды репликации, оболочка инъецированного ядра должна стать проницаемой

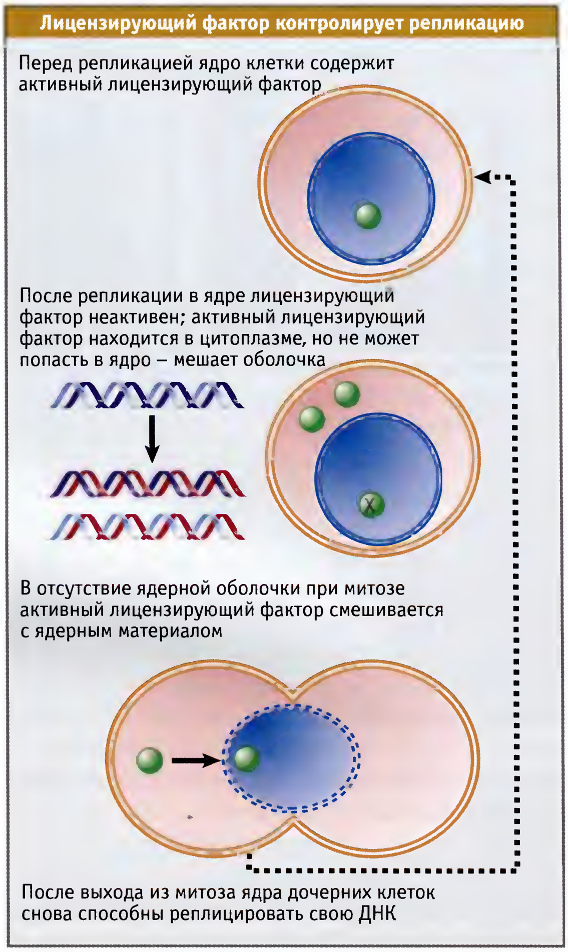


РИС. 15.14 Лицензирующий фактор в ядре инактивируется после репликации. Новая порция этого фактора может попасть в ядро не раньше, чем ядерная оболочка разрушится при митозе

Та же модельная система помогла решить задачи следующего уровня. Из яйцеклеток *X. laevis* был получен ядерный экстракт, поддерживающий репликацию *in vitro*. Экстракт разделили на компоненты, среди которых удалось идентифицировать белок с искомыми свойствами.

На РИС. 15.14 приведена модель контроля повторной инициации репликации, где этому белку отводится роль **лицензирующего фактора** (licensing factor). Фактор присутствует в ядре до репликации. Один раунд репликации инактивирует (возможно, разрушает) фактор; следующий раунд репликации не может состояться до тех пор, пока ядро не будет снабжено новой порцией фактора. Попасть в ядро из цитоплазмы фактор может только во время митоза, когда оболочка ядра начинает рваться и деградировать. Эта система регуляции приводит к двум последствиям. Во-первых, расход всего фактора в первом раунде репликации, предотвращает опасность множественности раундов в одном клеточном цикле. Во-вторых, здесь прослеживается обратная связь деления с репликацией: инициация репликации возможна, если после прошлой репликации клетка прошла митоз.

15.9 Лицензирующий фактор состоит из белков МСМ

Ключевые понятия:

- У дрожжей белковый комплекс ORC ассоциирован с ориджинами репликации во всех фазах клеточного цикла.
- Нестабильный белок Cdc6 синтезируется только в фазе G1.
- Белок Cdc6, связываясь с ORC, разрешает белкам МСМ также связаться с ORC.
- После того как репликация инициирована, белки Cdc6 и МСМ уходят; деградация Cdc6 препятствует повторной инициации.
- Некоторые белки МСМ присутствуют в ядре всегда, остальные могут попасть в ядро только во время митоза.

Поведение комплексов ORC в ориджинах является ключевым моментом контроля репликации. Мы помним, что ORC — это белковый комплекс массой 400 кДа; у дрожжей *S. cerevisiae* он связывается с элементом *ARS* (см. разд. 15.7 «Выделение ориджинов из геномов дрожжей»). В элементе *ARS*, работающем как ориджин репликации, есть сайт А с четким консенсусом, а также три сайта В (см. рис. 15.12). ORC, комплекс из шести белков, каждый из которых, судя по фенотипам мутантов, очень важен, связывается одновременно с сайтами А и В1 (ближайшим из сайтов В). Для связывания необходим АТФ, который, впрочем, гидролизуется на более поздней стадии. Фактор транскрипции ABF1, в свою очередь, связывается с сайтом В3. ABF1 помогает инициации репликации, которую вызывают события на сайтах А и В1 с участием ORC.

Большинство ориджинов расположено между генами. Возможно, для инициации репликации важно, чтобы в окрестном хроматине не шла никакая транскрипция. Мы помним, что ORC свя̀зан с ориджином на протяжении всего клеточного цикла. Однако, в зависимости от взаимодействия других белков с комплексом ORC–ориджин, характер связывания ORC с ДНК меняется. Изменения в «паттерне защиты от нуклеаз», который ORC ориентирует на ДНК, отражают циклический ход событий в ориджине. Представить себе эти события можно, взглянув на РИС. 15.15.

В конце клеточного цикла (то есть перед делением клетки) ORC связан с сайтами А и В1. При этом «паттерн защиты» похож на аналогичный паттерн, возникающий при связывании ORC со свободной ДНК *in vitro*. Защищенными оказываются и А, и В1, и промежуток между ними, однако посередине В1 есть очень короткий сайт, гиперчувствительный к ДНКазе. В фазе G1 этот паттерн меняется. Самое явное изменение, утрата гиперчувствительного сайта, происходит из-за того, что с ORC связывается белок Cdc6. У дрожжей белок Cdc6 очень нестабилен, с периодом полураспада менее 5 мин. Он синтезируется в фазе G1 и ассоциирован с ORC от выхода из митоза до поздней G1. Из-за того что Cdc6 быстро деградирует, на более поздних стадиях клеточного цикла его уже нет и в помине. (В клетках млекопитающих Cdc6 более стабилен, так что механизм его отстранения иной: во время S-фазы он фосфорилируется и в таком виде подлежит экспорту из ядра.) У Cdc6 есть А
Фазная активность, необходимая для его функции. С
служит адаптором между ORC и еще одним белковь
комплексом, участвующим в лицензировании и иниц
ации репликации.

У дрожжей обнаружено несколько мутаций, р
бота с которыми помогла понять, как устроена си
тема контроля доступа лицензирующего фактор
По идее, мутанты, у которых не работает собствен
лицензирующий фактор, не должны инициирова
репликацию. Именно так проявляют себя мутации
генах, кодирующих белки МСМ2, МСМ3 и МСМ5 (
англ. minichromosome maintenance — поддержан
минихромосомы). Другие мутации, бьющие по сист
ме, инактивирующей лицензирующий фактор (пос.
того, как репликация стартовала), должны приводи
к накоплению избыточных количеств ДНК, посколь
продленное присутствие лицензирующего факто
должно, по идее, создавать возможность для повто
ной репликации. Мутации с таким эффектом соср
доточены в генах, кодирующих компоненты систем
убиквитинирования, ответственной за деградаци
определенных белков. Скорее всего, утилизация л

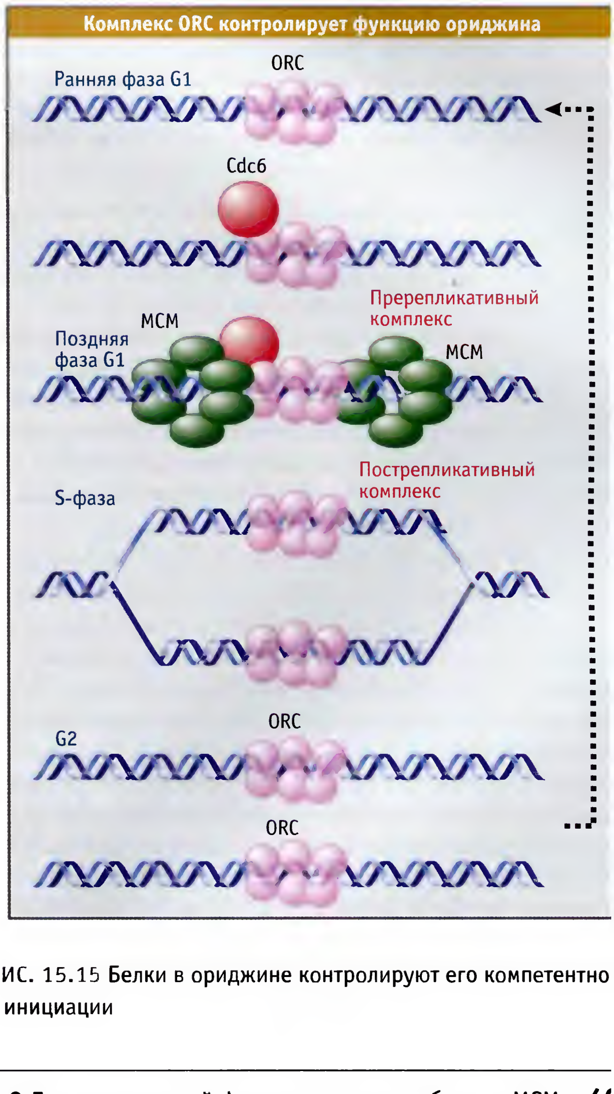


РИС. 15.15 Белки в ориджине контролируют его компетентно
к инициации

цензирующего фактора производится при участии именно этой системы.

Белки MCM2,3,5 необходимы для репликации. Попасть в ядро они могут только во время митоза. В клетках животных есть их гомологи; MCM3 там связан с нереплицированным хромосомным материалом и высвобождается после репликации. У животных комплекс MCM2,3,5 весь клеточный цикл проводит в ядре; из-за этого он может считаться лишь одним из компонентов лицензирующего фактора. Другой компонент лицензирующего фактора, способный проникнуть в ядро только во время митоза, может быть необходим для ассоциации MCM2,3,5 с хромосомным материалом.

У дрожжей присутствие в ориджине белка Cdc6 позволяет белкам MCM присоединиться к комплексу инициации репликации. Таким образом, ориджин вступает в S-фазу в виде **предрепликативного комплекса**. В нем состоят ORC, Cdc6 и белки MCM. После того как произошла инициация, Cdc6 и MCM уходят, приводя ориджин к состоянию **пострепликативного комплекса**. Из белков в нем состоит один только ORC. В S-фазе Cdc6 быстро деградирует и не может поддержать перезагрузку белков MCM в ориджин; из-за этого ориджин не может быть использован повторно.

Если при помощи техник эктопической экспрессии наладить поступление Cdc6 в ориджин в фазе G2, то белки MCM все равно не будут активированы до тех пор, пока не наступит фаза G1. Это говорит о том, что существует дополнительный механизм предохранения ориджинов от несвоевременного связывания с MCM. Этот механизм можно считать замыкающим контактом в регуляторном контуре лицензирования инициации репликации. Контроль лицензирования у дрожжей осуществляется не на уровне доступа фактора в ядро; в этом заключается различие между ними и клетками животных.

У *S. cerevisiae* белки MCM2-7 образуют шестичленный комплекс, замкнутый в кольцо вокруг ДНК. Некоторые белки из комплекса ORC имеют сходство с белками собственно репликации, наводящими ДНК-полимеразу на ДНК. Возможно, комплекс ORC, используя гидролиз АТФ, примерно так же накладывает на ДНК кольцо из MCM.

В экстрактах клеточных ядер *X. laevis* репликация может быть инициирована, даже если удалить ORC после того, как он осуществил доставку белков Cdc6 и MCM. Таким образом, роль ORC можно сравнить с ролью инструктора, обязанного показать ориджин белкам Cdc6 и MCM, контролирующим, соответственно, инициацию и лицензирование.

Белки MCM в репликации нужны не только для инициации, но и для элонгации. После того как репликация стартовала, они продолжают работать в репликативной вилке. Их точная роль в элонгации не выяснена; возможно, вносят какой-то вклад в хеликазную активность при расплетании ДНК. А возможно, белки MCM участвуют в элонгации в качестве авангарда, подготавливающего хроматин к воздействию хеликазы на ДНК.

15.10 В митохондриях ориджины репликации имеют вид D-петель

Ключевые понятия:

- Для инициации репликации двух цепей мтДНК используются отдельные ориджины с разными последовательностями.
- Репликация H-цепи инициируется в D-петле.
- Репликация L-цепи инициируется, когда первая репликативная вилка достигнет ориджина.

У всех организмов хромосомные ориджины используются для инициации репликации в конкретно заданное время. Они состоят из последовательностей ДНК, узнаваемых функциональными системами прямо в двухцепочечной форме. Инициация репликации требует разделения комплементарных цепей ДНК с последующим началом синтеза новых цепей в обоих направлениях. В митохондриях, с их собственной ДНК, все устроено по-другому.

Репликация кольцевой двухцепочечной молекулы мтДНК тоже стартует во вполне конкретном ориджине со специфической последовательностью. Но сначала в качестве матрицы для синтеза дочерней цепи используется только одна из цепей исходного дуплекса (когда речь идет о мтДНК млекопитающих, эту цепь называют H-цепью). Полимераза пробегает небольшую дистанцию, смещая с H-цепи ее партнера, L-цепь исходного дуплекса; это изображено на РИС. 15.16. При этом L-цепь становится одиночной. Такой участок выглядит очень характерно и его называют D-петлей (от *англ.* **D loop**, displacement loop — петля смещения).

ДНК-полимеразы не начинают синтез ДНК без затравочного 3′-конца (см. разд. 18.8 «Чтобы начать синтез ДНК, необходима затравка»). В ориджине H-цепи репликация начинается после того, как РНК-полимераза синтезирует там РНК (которая будет продолжать оставаться в гибриде со своей ДНК-матрицей). Эта РНК разрезается по нескольким дискретным сайтам. В каждый разрез открывается свободный 3′-OH конец, с которого ДНК-полимераза в принципе может начать синтез новой цепи ДНК. Эндонуклеаза, вносящая в РНК эти разрезы, обладает специфичностью к тройственной структуре: гибрид ДНК-РНК в сопровождении смещенной одиночной цепи ДНК с идентичной последовательностью. После того как свободные 3′-OH концы подготовлены, ДНК-полимераза использует один из них как затравку для синтеза цепи ДНК.

В мтДНК млекопитающих образуется одна D-петля, имеющая собственно петлю протяженностью до 500–600 пн. Короткая новосинтезированная L-цепь ДНК испытывает в этой D-петле большую нагрузку: именно ей приходится поддерживать исходный дуплекс в раскрытом состоянии. Сначала эта цепь много раз деградирует, и ее приходится синтезировать заново, пока наконец она не достигнет такого размера, что синтез ДНК на матрице L-цепи тоже сможет продолжиться.

В мтДНК других организмов можно насчитать несколько D-петель (из-за того что используется сразу

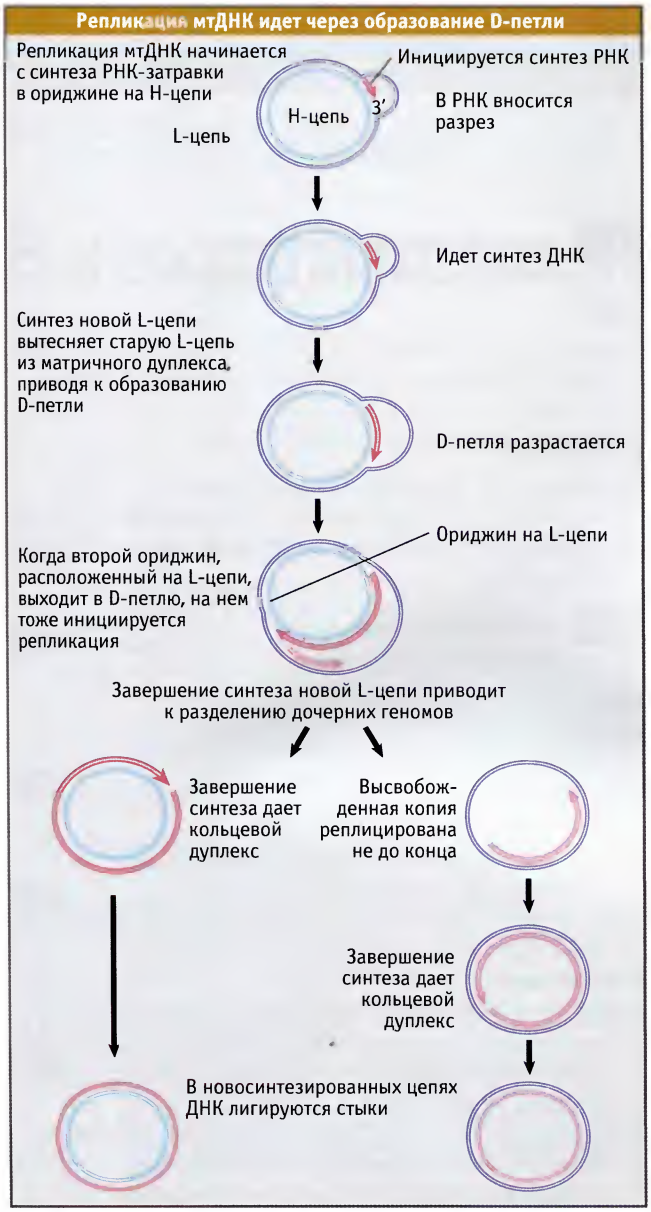


Рис. 15.16 **В митохондриях млекопитающих каждая цепь кольцевого дуплекса ДНК имеет свой собственный ориджин репликации. Во время репликации митохондриальной ДНК раскрытие матричного дуплекса поддерживается D-петлей**

несколько ориджинов). Репликация ДНК хлоропластов инициируется точно так же; в ДНК хлоропластов высших растений образуется две D-петли.

Итак, по мере того как D-петля разрастается, смещенный участок исходной L-цепи становится все длиннее. Так продолжается до тех пор, пока экспансия не пройдет примерно две трети круга, достигнув участка, где находится ориджин смещенной L-цепи. Там будет инициирован синтез новой Н-цепи: специальная праймаза узнает на смещенной L-цепи свой сайт и синтезирует короткую РНК-затравку. ДНК-полимераза начнет быстро наращивать эту РНК по матрице смещенной L-цепи в направлении, противоположном синтезу новой L-цепи.

Из-за того что ее синтез стартует с таким опозданием, новая Н-цепь к моменту окончания синтеза L-цепи успевает синтезироваться лишь на треть. В этот момент высвобождается две дочерних кольцевых молекулы: одна из них представляет собой полноценный кольцевой дуплекс; другая имеет одноцепочечную брешь, которая заполняется по мере того, как синтез новой Н-цепи идет к завершению. Наконец, в каждом из кольцевых дочерних дуплексов производится склеивание (лигирование) концов новосинтезированной цепи.

Инициация репликации через образование D-петель свидетельствует о принципиальном отличии митохондриальных ориджинов от хромосомных. С учетом сведений, изложенных в этом разделе, концепцию ориджина следует немного изменить, сказав, что *ориджин может быть областью ДНК, которая служит для инициации синтеза ДНК с использованием одной из цепей в качестве матрицы*. Раскрытие дуплекса ДНК не обязательно сопровождается инициацией синтеза ДНК на матрицах обеих цепей. В случае мтДНК ориджины репликации двух комплементарных цепей находятся на большом расстоянии друг от друга. Ориджины, которые существуют для репликации только одной из двух комплементарных цепей, характерны также для систем репликации по типу «катящегося кольца» (см. разд. 16.4 «С «катящихся колец» сходят мультимеры репликонов»).

15.11 Резюме

В каждом клеточном цикле хромосомная ДНК реплицируется единожды. Инициация репликации обрекает клетку на деление; завершение репликации может служить сигналом, запускающим собственно процесс деления. Бактериальная хромосома состоит из единственного репликона; эукариотическая хромосома поделена на множество репликонов, активность которых приходится на продолжительный интервал клеточного цикла, S-фазу. Проблема репликации краев линейного репликона может решаться по-разному; чаще всего она решается путем придания репликону формы кольца. У некоторых вирусов есть специальные белки, которые узнают концы ДНК. В эукариотических клетках проблема актуальна только для репликонов по краям хромосом.

Минимальный ориджин *E. coli* имеет длину ~245 пн; в нем инициируется двунаправленная репликация. Две репликативные вилки выходят из ориджина и обходят хромосому с разных сторон, до тех пор, пока не встретятся. Если встреча репликативных вилок друг с другом не остановила их движения, дальше по ходу их ждут специально предусмотренные участки терминации, *ter*. В клетке *E. coli* может реплицироваться любая молекула ДНК, содержащая последовательность минимального ориджина. Транскрипционные домены у *E. coli* устроены таким образом, что обычно транскрипция идет в том же направлении, что и репликация.

У эукариот синтез ДНК происходит в десятки раз медленнее, чем у бактерий. Во время S-фазы ориджины работают на запуск двунаправленной репликации; вероятно, между ними соблюдается определенная очередность. В ориджинах *S. cerevisiae* обнаруживается консенсус из 11 пар оснований, A-T и T-A.

После деления эукариотической клетки ядра дочерних клеток содержат лицензирующий фактор, необходимый для инициации репликации. После инициации репликации этот фактор разрушается; из-за этого последующие раунды репликации не могут состояться. Лицензирующий фактор не может быть импортирован в ядро из цитоплазмы; у высших эукариот его присутствие в ядре может быть восстановлено только после образования разрывов в оболочке ядра во время митоза.

У дрожжей ориджины узнаются белковым комплексом ORC; связь поддерживается в продолжение всего клеточного цикла. Начиная с конца митоза белок Cdc6 присутствует в ядре; но к концу S-фазы его там уже нет. У дрожжей Cdc6 синтезируется в фазе G1 и быстро деградирует. В клетках млекопитающих синтез Cdc6 идет постоянно, но в S-фазе он подлежит экспорту из ядра. Присутствие Cdc6 позволяет белкам MCM связаться с ориджином. Белки MCM необходимы для инициации репликации. Функция лицензирования обеспечивается Cdc6 совместно с белками MCM.

Литература

15.1 Введение

Исследование

Jacob, F., Brenner, S., and Cuzin, F. (1963). On the regulation of DNA replication in bacteria. *Cold Spring Harbor Symp. Quant. Biol.* 28, 329–348.

15.2 Репликоны бывают линейными и кольцевыми

Обзор

Brewer, B. J. (1988). When polymerases collide: replication and transcriptional organization of the *E. coli* chromosome. *Cell* 53, 679–686.

Исследование

Cairns, J. (1963). The bacterial chromosome and its manner of replication as seen by autoradiography. *J. Mol. Biol.* 6, 208–213.

Iismaa, T. P. and Wake, R. G. (1987). The normal replication terminus of the *B. subtilis* chromosome, *terC*, is dispensable for vegetative growth and sporulation. *J. Mol. Biol.* 195, 299–310.

Liu, B., Wong, M. L., and Alberts, B. (1994). A transcribing RNA polymerase molecule survives DNA replication without aborting its growing RNA chain. *Proc. Natl. Acad. Sci. USA* 91, 10660–10664.

Steck, T. R. and Drlica, K. (1984). Bacterial chromosome segregation: evidence for DNA gyrase involvement in decatenation. *Cell* 36, 1081–1088.

Zyskind, J. W. and Smith, D. W. (1980). Nucleotide sequence of the *S. typhimurium* origin of DNA replication. *Proc. Natl. Acad. Sci. USA* 77, 2460–2464.

15.3 Ориджины можно картировать при помощи авторадиографии и электрофореза

Исследование

Huberman, J. and Riggs, A. D. (1968). On the mechanism of DNA replication in mammalian chromosomes. *J. Mol. Biol.* 32, 327–341.

15.6 Эукариотические хромосомы состоят из множества репликонов

Обзор

Fangman, W. L. and Brewer, B. J. (1991). Activation of replication origins within yeast chromosomes. *Annu. Rev. Cell Biol.* 7, 375–402.

Исследование

Blumenthal, A. B., Kriegstein, H. J., and Hogness, D. S. (1974). The units of DNA replication in *D. melanogaster* chromosomes. *Cold Spring Harbor Symp. Quant. Biol.* 38, 205–223.

15.7 Выделение ориджинов из геномов дрожжей

Обзор

Bell, S. P. and Dutta, A. (2002). DNA replication in eukaryotic cells. *Annu. Rev. Biochem.* 71, 333–374.

DePamphlis, M. L. (1993). Eukaryotic DNA replication: anatomy of an origin. *Annu. Rev. Biochem.* 62, 29–63.

Gilbert, D. M. (2001). Making sense of eukaryotic DNA replication origins. *Science* 294, 96–100.

Kelly, T. J. and Brown, G. W. (2000). Regulation of chromosome replication. *Annu. Rev. Biochem.* 69, 829–880.

Исследование

Anglana, M., Apiou, F., Bensimon, A., and Debatisse, M. (2003). Dynamics of DNA replication in mammalian somatic cells: nucleotide pool modulates origin choice and interorigin spacing. *Cell* 114, 385–394.

Chesnokov, I., Remus, D., and Botchan, M. (2001). Functional analysis of mutant and wild-type *Drosophila* origin recognition complex. *Proc. Natl. Acad. Sci. USA* 98, 11997–12002.

Marahrens, Y. and Stillman, B. (1992). A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science* 255, 817–823.

Wyrick, J. J., Aparicio, J. G., Chen, T., Barnett, J. D., Jennings, E. G., Young, R. A., Bell, S. P., and Aparicio, O. M. (2001). Genome-wide distribution of ORC and MCM proteins in *S. cerevisiae*: high-resolution mapping of replication origins. *Science* 294, 2357–2360.

15.10 В митохондриях ориджины репликации имеют вид D-петель

Обзор

Clayton, D. (1982). Replication of animal mitochondrial DNA. *Cell* 28, 693–705.

Clayton, D., A. (1991). Replication and transcription of vertebrate mitochondrial DNA. *Annu. Rev. Cell Biol.* 7, 453–478.

Shadel, G. S. and Clayton, D. A. (1997). Mitochondrial DNA maintenance in vertebrates. *Annu. Rev. Biochem.* 66, 409–435.

Внехромосомные репликоны

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

16.1 Введение

16.2 Края линейной ДНК сложно реплицировать

- Чтобы реплицировать цепь ДНК со свободным 5′-концом, нужны специальные ухищрения.

16.3 Инициация на концах вирусной ДНК возможна благодаря терминальным белкам

- Терминальный белок, связанный с 5′-концом матрицы, несет цитидин со свободной 3′-OH группой, который служит затравкой для синтеза ДНК.

16.4 С «катящихся колец» сходят мультимеры репликонов

- Первичным продуктом репликации по типу «катящегося кольца» является одноцепочечный мультимер последовательности репликона.

16.5 Геномы некоторых фагов реплицируются по принципу «катящегося кольца»

- Белок А фага φX работает как *цис*-действующая релаксаза — генератор кольцевых одноцепочечных молекул из реплицированного материала, сходящего с «катящегося кольца».

16.6 Фактор F передается от бактерии к бактерии при конъюгации

- Фактор F может существовать автономно, как плазмида, поддерживаясь в числе одной копии на бактериальную хромосому.
- Фактор F может внедриться в бактериальную хромосому; в этом случае его собственные функции исполнения и контроля репликации будут подавлены.
- Фактор F кодирует систему образования особых ворсинок на поверхности бактерии.
- Ворсинка позволяет F-положительной бактерии вступать в контакт с F-отрицательной бактерией и инициировать конъюгацию.

16.7 При конъюгации происходит перенос одноцепочечной ДНК

- Перенос фактора F при конъюгации инициируется после того, как в *oriT* стартовала репликация по типу «катящегося кольца».
- Инициатором переноса выступает свободный 5′-конец одной из цепей ДНК фактора F.
- ДНК донора, перенесенная в цитоплазму реципиента, переводится в двухцепочечную форму.
- Если фактор F у донора является плазмидой, то при конъюгации реципиент получает одну копию этой плазмиды.
- Если фактор F у донора встроен в его собственный геном, то при конъюгации реципиент получает неполную копию хромосомы донора; перенос хромосомы занимает много времени и прерывается в произвольный момент в связи с окончанием акта конъюгации.

16.8 Плазмида Ti, населяющая бактериальные клетки, является возбудителем корончатого галла у растений

- Клетки растений, инфицированные бактерией *A. tumefaciens*, могут трансформироваться в опухолевые.
- Опухолеродная трансформация растительной клетки происходит из-за плазмиды, хранящейся в бактерии.
- Эта плазмида несет в себе гены синтеза и метаболизма опинов (производных аргинина); данные гены экспрессируются в опухолях, вызванных бактерией.

16.9 Плазмида Ti содержит гены, необходимые для вторжения Т-ДНК в геном растения

- Копия обширной области плазмиды Ti переносится в ядро растительной клетки.
- Гены *vir* плазмиды Ti нужны для осуществления переноса, однако находятся вне области, переносимой в растительную клетку.
- Экспрессия генов *vir* индуцибельна; она начинается под действием фенольных соединений, которые выделяет растение с поврежденной тканью.
- При связывании с веществом-индуктором мембранный белок VirA фосфорилирует сам себя по гистидину.
- VirA переносит фосфат на VirG, тем самым активируя его
- Связка VirA-VirG представляет собой «фосфогистидиновое реле», типичный для бактерий компонент сигнального пути.

16.10 Перенос Т-ДНК схож с конъюгацией бактерий

- Т-ДНК образуется путем подрезания дуплекса ДНК у правой границы Т-области.
- Во вскрытом дуплексе есть затравочный 3′-конец, с него начинается синтез новой цепи ДНК, сопровождающийся вытеснением одной из цепей дуплекса; вытесненный фрагмент одиночной цепи ДНК уходит из бактерии в ядро растительной клетки.
- Когда синтез новой цепи ДНК доходит до левой границы Т-области, перенос терминируется.
- Одноцепочечная Т-ДНК переносится в растительную клетку в комплексе с белком VirE2.
- После переноса Т-ДНК переходит в двухцепочечную форму и подлежит интеграции в геном растения.
- Несмотря на то что механизм интеграции изучен не полностью, Т-ДНК активно используется в лабораторной практике для доставки новых последовательностей в геном растительной клетки.

16.11 Резюме

16.1 Введение

Бактериальная клетка, обладая собственной хромосомой, может быть также реципиентом («хозяином») других генетических единиц, реплицирующихся независимо. Эти так называемые внехромосомные геномы принадлежат плазмидам и фагам. Плазмиды — это кольцевые молекулы ДНК, поддерживаемые клеткой в характерном числе копий. Однокопийные плазмиды поддерживаются в числе, равном числу хромосом бактерии-хозяина — одна на единичную клетку. Во время деления клетки, при поддержке специального молекулярного аппарата, дубликаты однокопийных плазмид распределяются в дочерние клетки — подобно дубликатам хромосомы. Многокопийные плазмиды, напротив, существуют во множестве копий на единичную клетку; при делении они могут распределяться по дочерним клеткам бесконтрольно, поскольку их так много, что каждой из дочерних клеток наверняка достанется какая-то часть плазмидных копий. Важное различие между плазмидами и фагами заключается в том, что плазмиды, при всей своей автономности в качестве генетических единиц, никогда не покидают цитоплазму бактерии. Фаги же являются вирусами бактерий: в конце инфекционного цикла они упаковывают свой геном (ДНК или РНК) в оболочку из белка и в таком виде выходят из бактерии наружу. Некоторые плазмиды и все фаги имеют свойство передаваться от бактерии-донора к бактерии-реципиенту; процесс заражения реципиента является обязательной фазой в инфекционном цикле любого фага.

Обязательным свойством (и отличительной особенностью) плазмиды или фага является способность существовать в бактериальной клетке в качестве независимой генетической единицы. Некоторые плазмиды и фаги могут также существовать в составе генома бактерии (как область бактериальной хромосомы, которая по своей нуклеотидной последовательности идентична геному плазмиды или фага) и наследоваться подобно собственным генам бактерии. Про фагов, которых можно обнаружить в составе бактериальной хромосомы, говорят, что они **лизогенные**. Плазмиды, способные так себя вести, называют **эписомами**. По механизмам встраивания в бактериальную хромосому и вырезания из нее лизогенные фаги и эписомы похожи друг на друга.

Между лизогенными фагами, плазмидами и эписомами можно провести еще одну, очень важную, параллель. Все эти репликоны производят впечатление «эгоистичных» по отношению к себе подобным: чем сильнее сходство, тем меньше шансов у двух репликонов утвердиться в одной бактериальной клетке. Такой эффект известен как **иммунность**. У плазмид иммунность является результатом работы систем контроля репликации; у лизогенных фагов его молекулярная природа другая.

На рис. 14.2 были показаны разные типы генетических единиц, способных распространять себя в бактериях в качестве независимых геномов. Так, у литического фага перенос геномов (в форме нуклеиновой кислоты) *между клетками* осуществляется через выпуск инфекционных частиц. Геном лизогенного фага обязательно представляет собой двухцепочечную ДНК, как и у плазмид и эписом. Некоторые плазмиды переносятся между клетками в процессе конъюгации, через установление прямого контакта между цитоплазмами донора и реципиента. При переносе по любому из двух типов плазмида или фаг могут случайно захватить с собой фрагмент генетического материала бактерии-хозяина. Поэтому перенос плазмид и фагов способствует обмену генетической информацией между хромосомами бактерий.

Важнейшая характеристика репликона — это манера использования ориджина. В бактериальных хромосомах и хромосомах эукариот ориджины используются для инициации единичных событий репликации, охватывающих весь репликон. Но репликоны могут быть объектами и совсем других событий. Самая потрясающая альтернатива происходит с небольшими независимыми единицами репликации некоторых вирусов. Как известно, в инфекционном цикле перед любым вирусом стоит задача наработать как можно большее число копий своего генома прежде, чем клетка-хозяин лопнет (лизируется), давая этим (специальным образом упакованным) копиям возможность заражать другие клетки. Так вот, некоторые вирусы реплицируются подобно тому, как это делает геном их хозяина: событие инициации ведет к продукции копий-дубликатов, после этого каждая копия проходит новый раунд репликации, и т. д. Но некоторые вирусы реплицируются иначе: одно событие инициации позволяет начитать огромный массив из множества копий генома, идущих стык в стык! Похожая программа может быть запущена эписомой, когда эта плазмидная ДНК, внедрившаяся в хромосому хозяина, перестает быть инертной в качестве репликона.

Многие репликоны из тех, что будут рассмотрены в этой главе, являются кольцевыми. На самом деле, для того чтобы репликация могла идти по типу продукции множества тандемных копий, репликон обязан быть замкнутым в кольцо. В отношении линейных внехромосомных репликонов наиболее интересным является решение проблемы репликации краев.

16.2 Края линейной ДНК сложно реплицировать

Ключевые понятия:

- Чтобы реплицировать цепь ДНК со свободным 5′-концом, нужны специальные ухищрения.

В гл. 15 «Репликон» мы уже познакомились с репликонами нескольких типов. Интересно, что ни один из них не имел свободного конца (и, соответственно, не имел концевой области). Эти репликоны либо были кольцевыми (как хромосома *E. coli* или мтДНК), либо входили в состав

более длинных единиц сегрегации (как репликоны в хромосомах эукариот). Однако линейные репликоны, у которых есть свободные концы, все же встречаются. Это могут быть одиночные внехромосомные единицы или репликоны, находящиеся на краях эукариотических хромосом.

То обстоятельство, что все полимеразы умеют вести синтез нуклеиновой кислоты только в одном направлении (удлиняя 3′-конец полимера), представляет собой, мягко говоря, проблему для синтеза ДНК на краях линейных репликонов. Чтобы это понять, рассмотрим цепи матричного дуплекса, изображенного на РИС. 16.1. У этого дуплекса есть один свободный конец. С репликацией нижней цепи никаких проблем не будет: полимераза «дочитает» ее до самого конца, после чего диссоциирует. С репликацией верхней цепи возникнет одна проблема: синтез должен стартовать с самого концевого основания матрицы. Если синтез стартует с отступом от конца матрицы — хотя бы на одно основание, — дочерняя цепь окажется укороченной. Это недопустимо, особенно если учесть, что общая потеря генетического материала будет прогрессировать с каждым раундом репликации.

Говорить об инициации репликации на самом конце линейной ДНК вообще-то бессмысленно. Сама посадка полимеразы на матрицу требует минимального расстояния: фермент должен окружить собой позицию, по которой в растущий полимер будет включено новое основание. Поэтому для репликации краев линейных репликонов должен быть в ходу какой-то совершенно особый механизм. Теоретически можно представить себе четыре варианта решения проблемы.

Для начала, проблемы можно избежать, превратив линейный репликон в кольцевую молекулу либо мультимер. Этот прием используется некоторыми фагами, например Т4 и лямбда (см. разд. 16.4 «С «катящихся колец» сходят мультимеры репликонов»).

На краю репликона ДНК может иметь нестандартную струтуру — например, складываться шпилькой, в которой спрятан открытый конец матрицы. Образование ковалентных сшивок по краям репликона используется при репликации линейной мтДНК инфузории *Paramecium*.

Вместо того чтобы быть жестко заданными, края репликона могут быть вариабельными. Проблема решается за счет возможности добавлять по краям репликона произвольные фрагменты, чтобы потом их терять. Именно этот способ применяется в хромосомах эукариот, где варьирует число коротких концевых повторов (см. разд. 28.18 «Теломеры синтезируются рибонуклеопротеиновым ферментом»).

Можно ли добиться того, чтобы синтез ДНК все-таки пошел с самого начала матрицы? В этом может помочь специальный белок. Такие белки есть у некоторых вирусов — они ковалентно пришиты к 5′-концевому основанию нуклеиновой кислоты. Такого рода ассоциация с белками свойственна аденовирусной ДНК, ДНК фага ϕ29, РНК полиовируса.

16.3 Инициация на концах вирусной ДНК возможна благодаря терминальным белкам

Ключевые понятия:

- **Терминальный белок, связанный с 5′-концом матрицы, несет** цитидин со свободной 3′-OH группой, **который служит затравкой для синтеза ДНК.**

Инициация репликации на концах линейной ДНК характерна для аденовирусов и вируса ϕ29. Репликация по принципу **смещения цепи** происходит в соответствии со схемой на РИС. 16.2. На одном из концов дуплекса инициируется синтез новой цепи; по мере своего роста эта цепь смещает параллельную цепь, которая исходно состояла в дуплексе с матричной цепью. Когда репликативная вилка доходит до другого конца молекулы, смещенная цепь становится свободной одноцепочечной молекулой. Она реплицируется независимо: по краям молекулы находятся короткие взаимно-комплементарные последовательности — инвертированные концевые

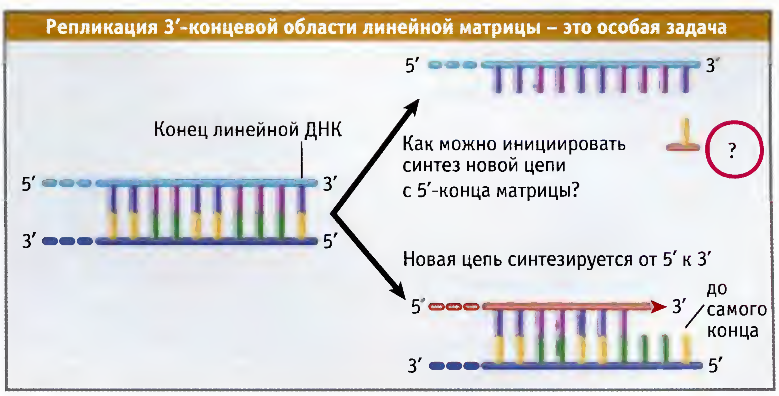


РИС. 16.1 Аппарат репликации умеет дочитывать линейную матрицу от 3′-конца. Может ли он инициировать репликацию 5′-конца?

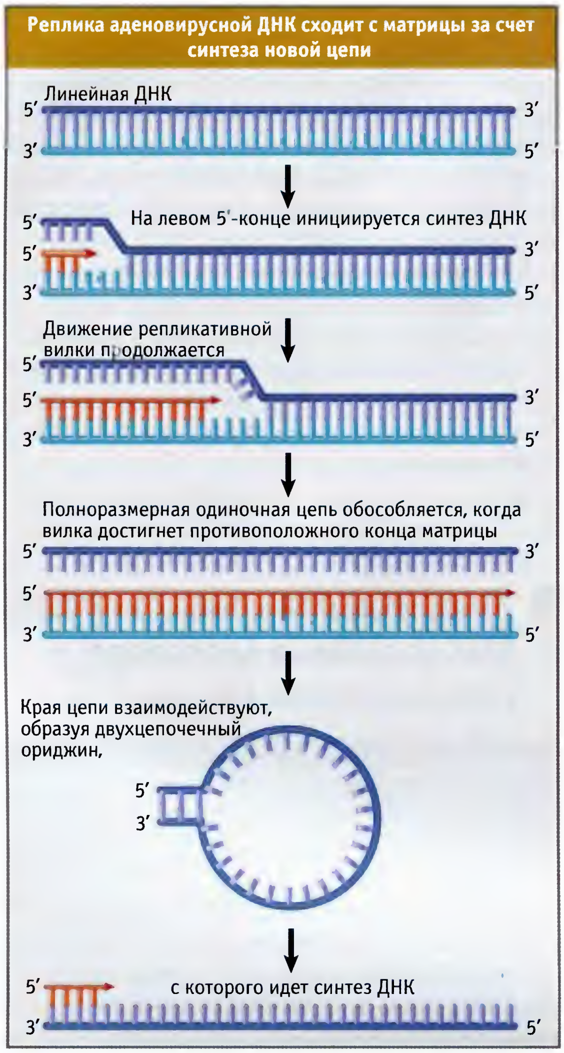


РИС. 16.2 Репликация каждой из цепей аденовирусной ДНК инициируется по-своему. Принципиальной особенностью реакции является вытеснение с матрицы одноцепочечной реплики за счет синтеза новой цепи

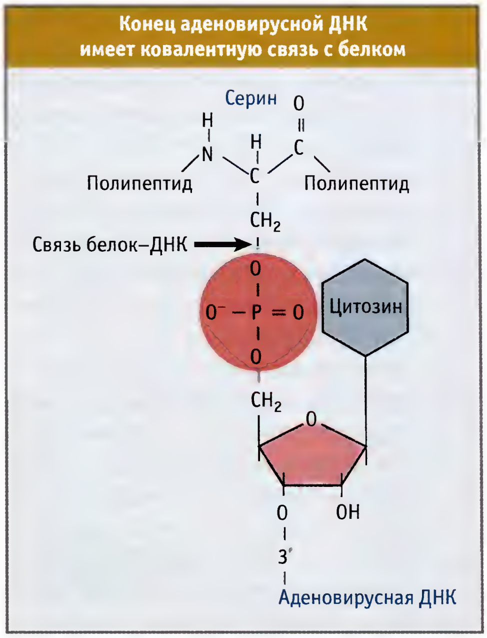


РИС. 16.3 5′-концевой фосфат двухцепочечной аденовирусной ДНК имеет ковалентную связь с серином в аденовирус-связывающем белке (55 кДа)

повторы. Они «залипают» друг на друга, образуя двухцепочечный ориджин.

У вирусов, которые реплицируются по принципу смещения цепи, есть специальный белок. После репликации этот белок оказывается ковалентно пришитым к 5′-концу цепи ДНК. У аденовирусов такой **терминальный белок** пришит к зрелой вирусной ДНК фосфодиэфирной связью между серином и 5′-концевым фосфатом ДНК. Это изображено на РИС. 16.3.

Каким образом присутствие терминального белка решает проблему инициации? Белок играет двойную роль: он и носитель цитидина со свободным 3′-OH концом, и адаптор для ДНК-полимеразы, с которой взаимодействует. Цитидин, которому предстоит стать затравкой в синтезе новой цепи, ДНК-полимераза сажает на терминальный белок сама — в присутствии аденовирусной ДНК. Репликация инициируется в соответствии со схемой на РИС. 16.4. Комплекс ДНК-полимеразы с терминальным

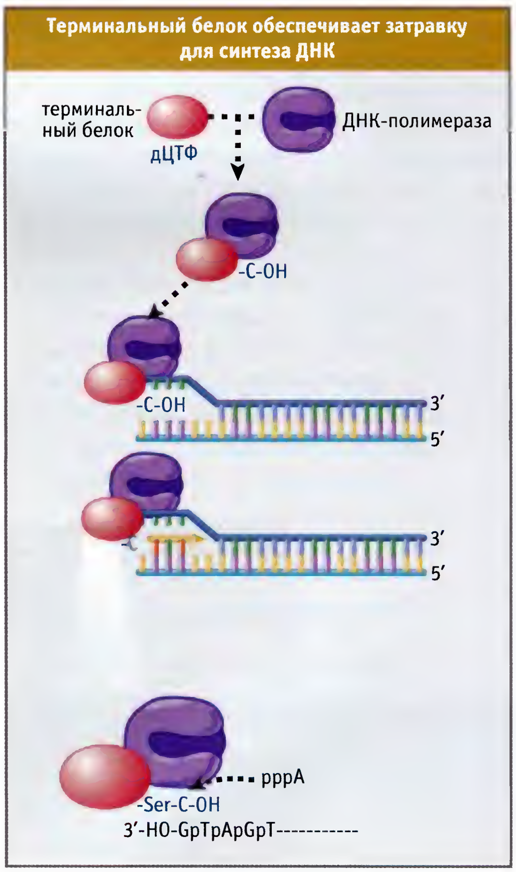


РИС. 16.4 Терминальный белок аденовируса связывается с 5′-концом ДНК и предоставляет свободный C-OH-конец, который служит затравкой для синтеза ДНК

белком, несущим затравочный цитидин, связывается с концом аденовирусной ДНК. После этого свободная 3'-OH группа цитидина используется ДНК-полимеразой как затравка для реакции элонгации. От этой реакции получается новая цепь ДНК, чей 5'-концевой нуклеотид, цитидин, пришит к серину терминального белка.

Реакция взаимодействия терминального белка с ДНК носит характер замещения. Дело в том, что в момент инициации на 5'-конце цепи, которую белку надо вытеснить из дуплекса, находится терминальный белок, который использовался в предыдущем раунде репликации. До инициации нового раунда старый терминальный белок продолжает занимать свой сайт на ДНК; во время инициации происходит его замещение.

Терминальный белок взаимодействует с сайтом ДНК, отстоящим от конца на 9–18 пн. Смежный сайт, между позициями 17 и 48, служит для связывания клеточного белка NF1 (от *англ.* nuclear factor 1 — ядерный фактор 1), который тоже нужен для инициации. Таким образом, белковый комплекс инициации, по-видимому, формируется между позициями 9 и 48, на фиксированном расстоянии от позиции 0 — реального конца ДНК.

16.4 С «катящихся колец» сходят мультимеры репликонов

Ключевые понятия:

- Первичным продуктом репликации по типу «катящегося кольца» является одноцепочечный мультимер последовательности репликона.

Какие именно структуры будет порождать репликация, зависит от свойств матрицы в сочетании с поведением репликативных вилок. Ключевыми факторами являются форма матрицы (кольцевая или линейная) и участие репликативной вилки в синтезе обеих комплементарных цепей ДНК либо только одной из них.

Репликация одной цепи используется для наработки копий некоторых кольцевых молекул. Сначала кольцевой дуплекс вскрывается надрезом одной из цепей. В надрез «смотрят» два конца ДНК, один из которых несет 3'-OH группу. ДНК-полимераза удлиняет этот свободный 3'-OH конец, одновременно вытесняя из дуплекса цепь, комплементарную матрице. Порядок событий изображен на РИС. 16.5.

Реплицирующуюся ДНК с такой структурой называют **катящимся кольцом**. Репликативная вилка в «катящемся кольце» совершает движение по кругу матричной цепи. В принципе она может продолжать описывать эти круги до бесконечности. По мере своего движения репликативная вилка удлиняет одну из цепей кольцевого дуплекса, вытесняя из дуплекса точно такую же цепь, синтезированную кругом раньше. Результат можно видеть в электронный микроскоп, фотография представлена на РИС. 16.6.

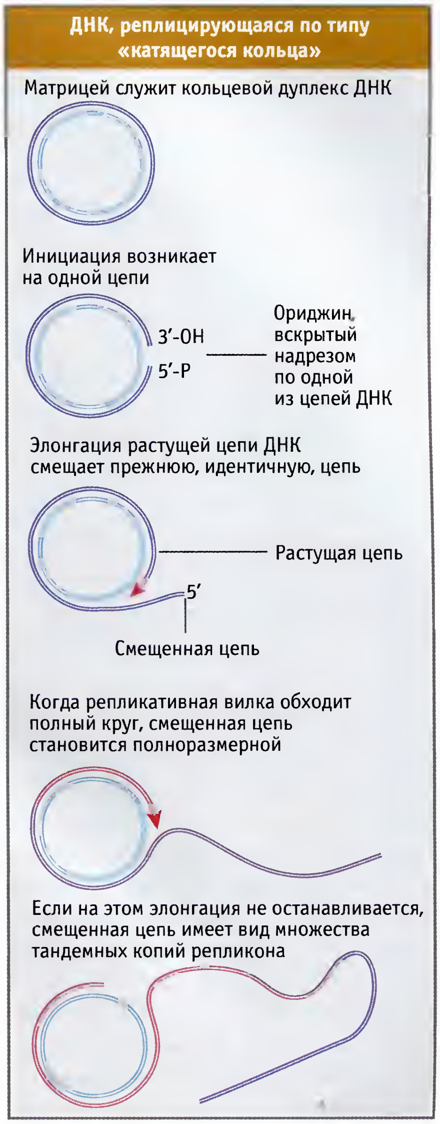


РИС. 16.5 «Катящееся кольцо» с мультимерным одноцепочечным хвостом



РИС. 16.6 В электронный микроскоп «катящееся кольцо» выглядит как кольцевая молекула с линейным хвостом. Фотография любезно предоставлена Ross B. Inman, Institute of Molecular Virology, Bock Laboratory and Department of Biochemistry, University of Wisconsin, Madison, Wisconsin, USA

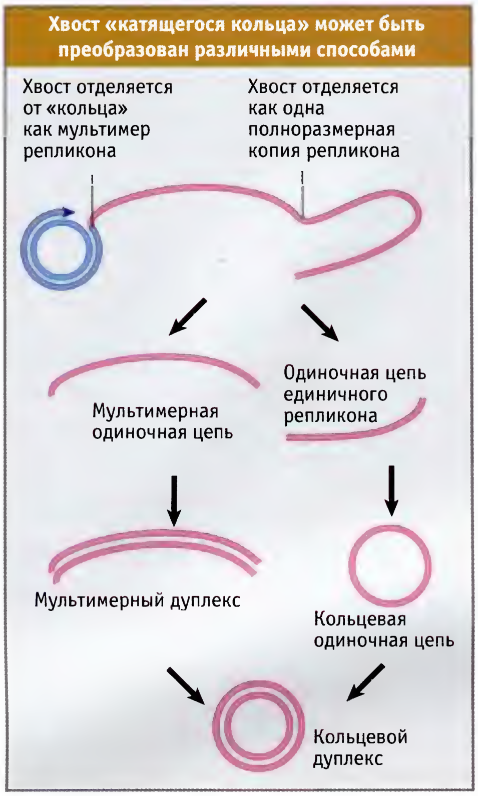


РИС. 16.7 Судьба хвоста определяет типы продуктов репликации. В одних случаях отделение от «кольца» хвостов единичной длины порождает одноцепочечные мономеры репликона; они могут быть переведены в двухцепочечную, кольцевую и двухцепочечную кольцевую формы. В других случаях от «кольца» отделяется хвост-мультимер; он представляет собой серию тандемно повторяющихся копий репликона. Заметим, что перевод хвоста в двухцепочечную форму может произойти еще до его отделения от «кольца»

Продукт такой репликации можно представить во времени (потраченном на его синтез) как пружину. Каждый новый виток пружины (копия «наружной» цепи дуплекса) образуется, когда репликативная вилка проходит по матрице свой очередной круг. Все витки пружины взаимнооднозначно проецируются на ее основание — «наружную» цепь исходного дуплекса, копией которого они являются.

Распрямив пружину, получим одноцепочечную молекулу. Начиная с ее 5′-конца идет последовательность единичного генома («наружной» цепи исходного дуплекса, бывшего основания пружины), за которой следуют подряд ее копии. Этих копий может быть сколько угодно. Так выглядит продукт репликации по типу «катящегося кольца».

Варианты естественной судьбы продукта репликации по типу «катящегося кольца» изображены на РИС. 16.7. Если его разрезать на единичные фрагменты, получится множество одноцепочечных линейных копий исходного кольцевого репликона. Можно оставить их такими, можно преобразовать их в дуплексы, синтезировав каждой копии комплементарную цепь (идентичную «внутренней» цепи исходного кольцевого репликона).

Репликация по типу «катящегося кольца» — прекрасный способ наработать какую-нибудь последовательность. В яйцеклетках шпорцевых лягушек *Xenopus laevis* она используется для амплификации рибосомальной ДНК (рДНК). Гены рибосомальной РНК (рРНК) организованы в большое число повторов. Для амплификации берется один такой повтор и превращается в «катящееся кольцо». У него растет хвост — продукт репликации, состоящий из множества повторов. Этот хвост, прямо в неразрезанном виде, подлежит преобразованию в двухцепочечную ДНК. Затем двухцепочечный хвост отрезают от «катящегося кольца» (а оно тем временем продолжает работать, и у него сразу же начинает расти новый хвост). Концы отрезанного двухцепочечного хвоста соединяют, получая большое кольцо рДНК. Это кольцо представляет собой амплифицированный материал: его нуклеотидная последовательность состоит из большого числа идентичных повторяющихся единиц.

16.5 Геномы некоторых фагов реплицируются по принципу «катящегося кольца»

Ключевые понятия:

- Белок А фага φX работает как *цис*-действующая релаксаза — генератор кольцевых одноцепочечных молекул из реплицированного материала, сходящего с «катящегося кольца».

Репликация ДНК, сопровождающаяся образованием «катящихся колец», свойственна многим бактериофагам. Одноцепочечный реплицированный материал можно нарезать в виде единичных геномов, которые затем можно использовать для последующих раундов репликации или упаковать в фаговые частицы. В этом разделе мы разберем один из частных случаев репликации фаговой ДНК по принципу «катящегося кольца». Нам поможет РИС. 16.8.

В начале инфекционного цикла геном бактериофага φX174 находится в форме одноцепочечной кольцевой ДНК; она обозначается как (+)-цепь. После того как она попала в бактериальную клетку и там дополнилась комплементарной (−)-цепью, молекула представляет собой кольцевой дуплекс, изображенный в верхней части рисунка. Кольцевому дуплексу специально придается сверхспирализованная форма. Белок А (закодированный в геноме фага) надрезает (+)-цепь дуплекса по конкретному сайту в ориджине репликации. После внесения надреза белок А сохраняет связь с 5′-концом (+)-цепи. Второй конец (+)-цепи (сво-

бодный 3′-конец) использует как затравку для синтеза ДНК ДНК-полимеразу.

В инициации по этой схеме очень важную роль играет структура ДНК. Дело в том, что белок А может надрезать кольцевой дуплекс, *только если этот дуплекс отрицательно сверхспирализован*, то есть если кольцо обвивает само себя в пространстве, причем направление вращения сверхвитков обратно направлению вращения витков двойной спирали ДНК (см. разд. 19.12 «Сверхспирализация затрагивает структуру ДНК»). Чтобы внести надрез, белок А должен связать одну из цепей дуплекса по участку в 10 пн, на котором находится сайт надреза. Вероятно, отрицательная сверхспирализация облегчает образование одноцепочечного участка, необходимого для посадки белка А на ДНК.

Итак, белок А надрезает дуплекс по одной из цепей и прикрепляется к свободному 5′-концу. Про такую энзиматическую активность говорят, что белок работает как **релаксаза**. В получившийся надрез смотрят два конца (+)-цепи ДНК: 3′-ОН конец (свободный) и 5′-фосфатный конец (пришитый к белку А ковалентной связью). Каждый из концов играет в репликации фX174 свою особую роль.

После того как надрез внесен, бактериальная ДНК-полимераза начинает наращивать свободный 3′-конец (+)-цепи на матрице кольцевой (–)-цепи: «кольцо» фX174 начинает «катиться». Полимераза обходит полный круг и возвращается в ориджин. Синтез продолжается, и вскоре с «катящегося кольца» сходит последовательность ориджина. Эту последовательность немедленно узнает белок А, который пришит ковалентной связью к смещенному 5′-концу той же (+)-цепи, а значит, находится на расстоянии одной длины генома фX174, если считать это расстояние по ДНК. Белок А снова проявляет себя как релаксаза: вносит надрез в последовательность ориджина и пришивается уже к новому, только что созданному им 5′-концу. В результате белок А продолжает оставаться при «катящемся кольце», а копия (+)-цепи, смещенная с «кольца» в первом раунде репликации, уходит со свободными концами. Этот циклический процесс может повторяться до бесконечности.

На самом деле, смещенная копия (+)-цепи уходит не со свободными концами, а в виде кольцевой одноцепочечной молекулы. Этому также содействует белок А. Переход белка А к инициации следующего раунда репликации сопряжен с соединением свободных концов единичной копии (+)-цепи.

У белка А есть одна интересная особенность: он является *цис*-действующим *in vivo*. (*Эта фраза означает, что белок А, синтезированный с транскрипта конкретной молекулы ДНК, будет способствовать репликации только этой конкретной молекулы ДНК.*) Чем это обусловлено, мы не знаем; в бесклеточной системе *in vitro* эффект не воспроизводится — там белок А готов работать с любой подходящей молекулой ДНК.

Во всяком случае, благодаря исследованиям активности белка А *in vitro*, можно представить себе, как этому белку удается оставаться ассоциированным с



Рис. 16.8 Репликативная вилка ДНК фX174 служит матрицей для синтеза одноцепочечных кольцевых копий генома этого фага. В каждом раунде репликации кольцевую матрицу сопровождает одна и та же молекула белка А. В конце каждого раунда белок А надрезает ориджин по (+)-цепи и перебрасывает себя на образовавшийся 5′-конец. Таким способом от «катящегося кольца» отделяется очередная (+)-цепь единичной длины; она сразу же становится кольцевой молекулой ДНК

одной и той же матричной (–)-цепью. У белка А есть два активных центра. Именно это позволяет ему разрезать новый ориджин, все еще сохраняя связь с предыдущим.

После того как смещенная копия (+)-цепи замкнулась в кольцо (при участии того же белка А), ее судьба зависит от текущей фазы инфекционного цикла бактериофага. Если это фаза репликации, кольцевая (+)-цепь будет дополнена (–)-цепью; получившийся дуплекс будет превращен в «катящееся кольцо» для наработки все новых копий генома бактериофага. Если инфекция в своем развитии достигла фазы формирования частиц, кольцевая (+)-цепь останется одиночной и будет упакована в инфекционную частицу бактериофага — для выхода во внешнюю среду после лизиса обреченной клетки.

16.6 Фактор F передается от бактерии к бактерии при конъюгации

Ключевые понятия:

- Фактор F может существовать автономно, как плазмида, поддерживаясь в числе одной копии на бактериальную хромосому.
- Фактор F может внедриться в бактериальную хромосому; в этом случае его собственные функции исполнения и контроля репликации будут подавлены.
- Фактор F кодирует систему образования особых ворсинок на поверхности бактерии.
- Ворсинка позволяет F-положительной бактерии вступать в контакт с F-отрицательной бактерией и инициировать конъюгацию.

Конъюгация у бактерий служит примером того, насколько тесно могут быть связаны процессы репликации автономной генетической единицы и ее распространения в популяции. При конъюгации геном плазмиды или фрагмент хромосомы бактерии-хозяина могут передаваться от одной бактерии к другой.

Конъюгация у бактерий *E. coli* опосредована **фактором F**. Он представляет собой классический пример эписомы — генетической единицы, способной существовать не только как свободная кольцевая плазмида, но и внедряться в бактериальную хромосому как линейная последовательность (подобно лизогенному бактериофагу). Плазмида F — это большая кольцевая ДНК (размером ~100 тпн). В бактериальной хромосоме есть несколько потенциальных сайтов интеграции фактора F. Чаще всего внедрение обусловлено рекомбинацией между определенными элементами ДНК, которые есть и в плазмиде F, и в хромосоме ее хозяина (их называют IS-элементами; подробнее см. в разд. 21.5 «Транспозоны перестраивают ДНК»). Плазмида F, находясь в свободной (неинтегрированной форме), использует свой собственный ориджин репликации, *oriV*, а также свою собственную систему контроля числа копий. Она поддерживается в виде единственной копии на клетку. После интеграции в бактериальную хромосому работа собственных систем репликации плазмиды F подлежит репрессии — ее ДНК теперь реплицируется как часть бактериальной хромосомы.

Представим себе популяцию бактерий, часть которой несет в себе фактор F (в свободной либо интегрированной форме). Эти F-положительные бактерии способны конъюгировать (спариваться) с F-отрицательными бактериями. При конъюгации формируется цитоплазматический мостик между F-положительной бактерией-донором и F-отрицательной бактерией-реципиентом; мостик используется для переноса фактора F. Если фактор F является плазмидой, он перейдет в реципиента как плазмида, и реципиент превратится из F-отрицательного в F-положительного. Если фактор F интегрирован в хромосому донора, он должен будет отдать реципиенту часть своей хромосомы.

Вообще-то, конъюгацию способны стимулировать и другие плазмиды. Но поскольку фактор F открыли и изучили раньше других, его система конъюгации рассматривается всеми как прототип этого способа переноса генетической информации.

В ДНК фактора F есть обширная область (~33 тпн), необходимая для конъюгации. Ее называют **областью переноса**. Область переноса содержит ~40 генов; с их расположением можно ознакомиться по карте на рис. 16.9. Названия локусов начинаются на *tra* и *trb*. Большинство генов области переноса экспрессируется в составе гигантской транскрипционной единицы *traY-I* протяженностью 32 тпн. Два гена, *traM* и *traJ*, экспрессируются отдельно. Один из них, *traJ*, является регулятором, запускающим экспрессию и *traM*, и *traY-I*. На комплементарной цепи ДНК расположен другой регулятор, *finP*: с него транскрибируется короткая антисмысловая РНК, выключающая *traJ*. Чтобы *finP* был активен, должен экспрессироваться еще один ген, *finO*. Большинство белков, которые продуцирует фактор F, занято в изменениях свойств поверхности бактериальной клетки и поддержании контакта между спаривающимися бактериями. Непосредственное отношение к переносу ДНК имеют только четыре продукта транскрипционной единицы *traY-I* (эти белки кодируются четырьмя генами *tra*).

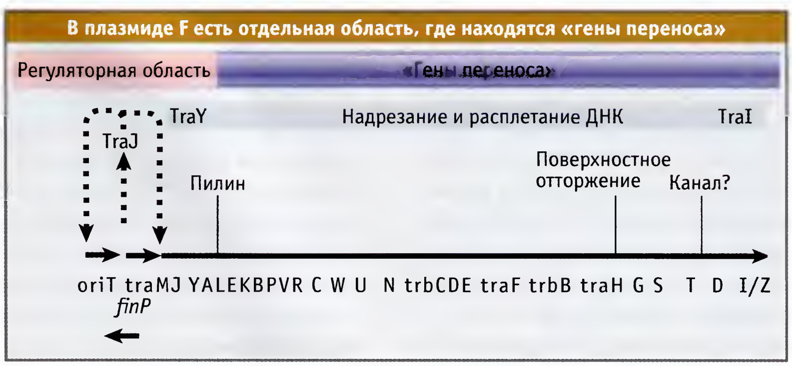


РИС. 16.9 Область *tra* плазмиды F содержит гены, необходимые для конъюгации бактерий

У F-положительных бактерий есть специальные выросты — так называемые ворсинки, или F-пили. Образование ворсинок обусловлено присутствием в клетке фактора F. Ген *traA* кодирует белок **пилин**; полимер пилина и есть ворсинка. Чтобы модифицировать незрелый пилин и собрать из него ворсинку, необходима экспрессия как минимум двенадцати генов *tra*. Ворсинка имеет вид полого цилиндра с внешним диаметром ~8 нм и диаметром центральной полости ~2 нм. Ворсинки торчат на поверхности бактерии и имеют длину от 2 до 3 мкм. Типичная F-положительная клетка имеет всего две-три ворсинки.

Спаривание двух бактерий инициируется в момент, когда кончик ворсинки нащупает поверхность клетки-реципиента. На РИС. 16.10 представлена фотография клеток *E. coli*, начинающих спаривание. Примечательно, что F-положительная клетка-донор не вступает в контакт с другими F-положительными клетками. Этот процесс контролируется продуктами генов *traS* и *traT* — белками «отторжения поверхностей». Из-за них бактерия очень вяло реагирует на возможность контакта с другими бактериями-носителями фактора F. Таким образом, потенциальный донор фактора F не может начать спаривание, пока не найдет себе F-отрицательного партнера.

Наличие у бактерии ворсинок (F-пилей) имеет ряд побочных эффектов. Например, ворсинки создают оптимальные условия для заражения клетки некоторыми бактериофагами. Ворсинки сорбируют частицы этих фагов, что создает плацдарм для инфекции; поэтому F-отрицательные бактерии устойчивы к ней.

Ворсинки обеспечивают первичный контакт донора с реципиентом, но этот контакт легко рушится. В дальнейшем продуктам других генов *tra* удается его стабилизировать, и клетки начинают идти на сближение. После этого ворсинки, которые были нужны лишь для того, чтобы инициировать спаривание, втягиваются или демонтируются, чтобы не мешать вхождению клеток в тесный контакт. Контакт должен привести к образованию цитоплазматического мостика — канала, по которому ДНК сможет перетечь из одной клетки в другую. У F-положительных бактерий есть белок внутренней мембраны TraD, он совершенно необходим для переноса ДНК. Возможно, TraD участвует в образовании канала.



РИС. 16.10 Спаривание инициируется после того, как F-пили донора войдут в контакт с поверхностью реципиента. Фотография любезно предоставлена Ron Skurray, School of Biological Sciences, University of Sydney.

16.7 При конъюгации происходит перенос одноцепочечной ДНК

Ключевые понятия:

- Перенос фактора F при конъюгации инициируется после того, как в *oriT* стартовала репликация по типу «катящегося кольца».
- Инициатором переноса выступает свободный 5'-конец одной из цепей ДНК фактора F.
- ДНК донора, перенесенная в цитоплазму реципиента, переводится в двухцепочечную форму.
- Если фактор F у донора является плазмидой, то при конъюгации реципиент получает одну копию этой плазмиды
- Если фактор F у донора встроен в его собственный геном, то при конъюгации реципиент получает неполную копию хромосомы донора; перенос хромосомы занимает много времени и прерывается в произвольный момент в связи с окончанием акта конъюгации.

Перенос фактора F инициируется в *oriT* (от *англ.* origin of transfer —ориджин переноса), который находится в его геноме у одной из границ области переноса. Инициация переноса становится возможна после того, как белок TraM подаст сигнал, что бактерии конъюгируют, и акт спаривания успел зайти достаточно далеко, чтобы начать перенос. Тогда другой белок, TraY, связывается с ДНК *oriT* и привлекает туда еще один белок, TraI, работающий релаксазой (как белок А фага φX174). Этот TraI вскрывает *oriT* в уникальном сайте *nic* и пришивает себя к образовавшемуся 5'-концу ДНК. Кроме того, TraI катализирует расплетание ~200 пн ДНК (это хеликазная активность, о ней см. в разд. 18.7 «Наработка одноцепочечной ДНК для размножения фага в рамках модельной системы φX»). Освободившись от TraI, лидирующий 5'-конец устремляется в бактерию-реципиента; это показано на РИС. 16.11. Синтез комплементарной цепи ДНК происходит уже в цитоплазме реципиента, который в итоге обретает статус F-положительного.

Донору нужно восполнить отсутствие ушедшей ДНК — синтезировать новую цепь на матрице оставшейся цепи. Если синтез новой цепи осуществляется одновременно с уходом переносимой цепи ДНК, плазмида F эквивалентна «катящемуся кольцу» с рис. 16.5 (но у нее нет протяженного одноцепочечного участка, изображенного на рис. 16.11). Действительно, ДНК, с которой происходит перенос цепи при конъюгации, обычно похожа на «катящееся кольцо». Однако репликация как таковая не является движущей силой переноса ДНК, то

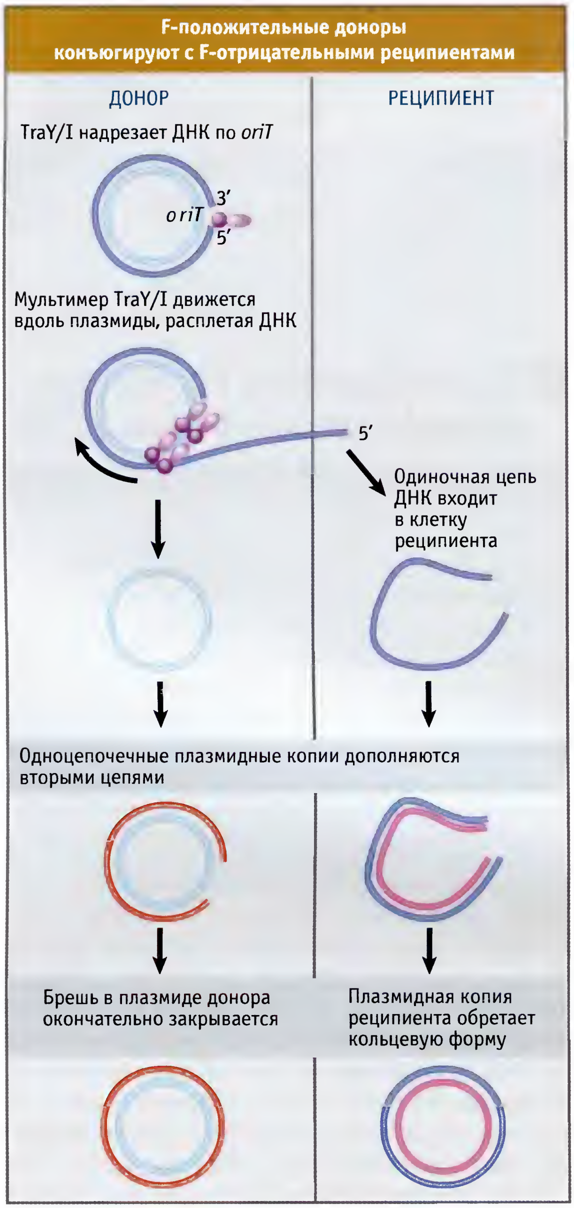


РИС. 16.11 **Перенос ДНК происходит после вскрытия плазмиды F по *oriT*. Лидирующий 5′-конец увлекает одиночную цепь ДНК в клетку реципиента; переносится всего одна полноразмерная копия плазмиды. В заключение, одноцепочечные плазмидные копии донора и реципиента «обрастают» комплементарными цепями**

есть перенос одноцепочечной молекулы не обусловлен синтезом ДНК. Бактерия-реципиент получает всего одну копию плазмиды F. Процесс смещения одиночной цепи с кольцевого дуплекса в цитоплазме донора терминируется после одного оборота, и целостность плазмиды F восстанавливается.

Теперь представим себе, что конъюгацию инициировал фактор F, который встроен в хромосому бактерии. Вскрытие *oriT* происходит всегда по одной и той же цепи ДНК, причем в такой ориентации, что лидирующий 5′-конец потащит за собой в реципиента лишь короткий край ДНК фактора F, за которым последует материал огромной хромосомы хозяина. Весь основной материал фактора F, с областью переноса, реципиенту не передается. Перенос одной из цепей огромной хромосомы будет идти до тех пор, пока не разорвется контакт между спаривающимися клетками (см. РИС. 16.12). Чтобы последовательность бактериальной хромосомы перенеслась в реципиента целиком, потребуется ~100 мин; в обычных условиях акт конъюгации менее продолжителен.

В цитоплазме бактерии-реципиента фрагмент геномной ДНК бактерии-донора переводится в двухцепочечную форму. Он может рекомбинировать с геномной ДНК хозяина; заметим,что для внедрения ДНК донора в хромосому реципиента требуется два события рекомбинации. Таким образом, конъюгация служит бактериям как средство обмена генетическим материалом — на

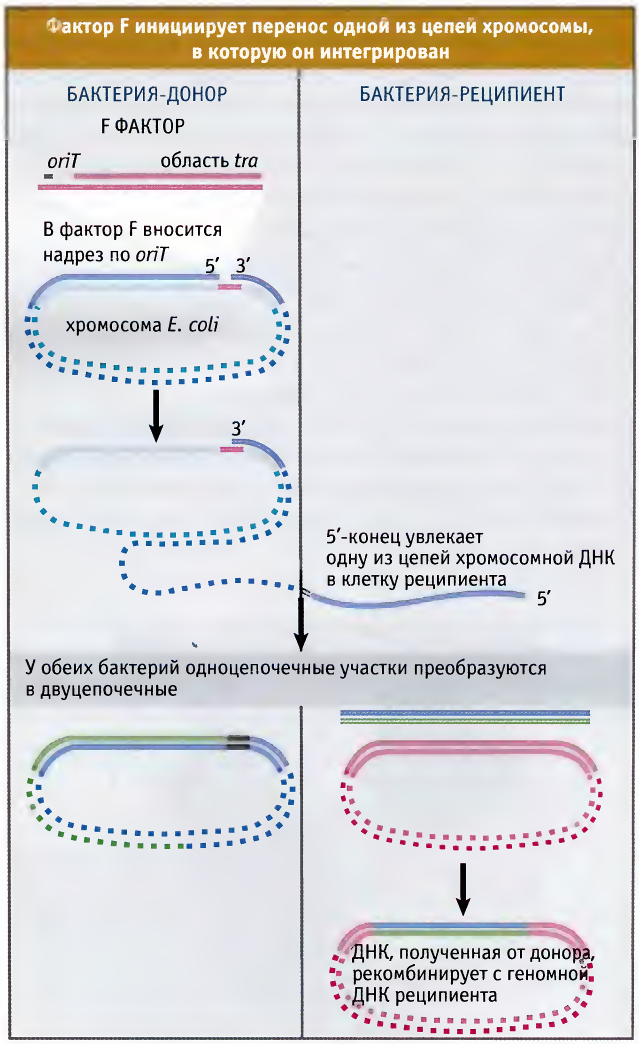


РИС. 16.12 **Перенос хромосомной ДНК происходит, когда фактор F, интегрированный в хромосому донора, надрезан по *oriT*. Переносимый генетический материал содержит лишь короткий фрагмент фактора F; за этим фрагментом следует материал хромосомы донора. Перенос продолжается до тех пор, пока конъюгирующие бактерии не выйдут из контакта друг с другом**

фоне их обычного бесполого роста в культуре. Штамм *E. coli*, где фактор F встроен в хромосому, способен навести в другом штамме, с которым он конъюгирует, высокую частоту появления рекомбинантных хромосом. Штаммы-доноры, от которых другие штаммы получают фрагменты хромосом, обозначаются как **Hfr** (от *англ.* high frequency recombination — рекомбинация с высокой частотой). После контакта с таким штаммом у других бактерий проявляется характерный паттерн новообретенных генетических маркеров, зависящий от положения, которое занимает фактор F в хромосоме Hfr.

Как уже говорилось, контакт между конъюгирующими бактериями обычно рвется прежде, чем одна из цепей хромосомы донора успеет перейти в реципиента целиком. Вероятность того, что конкретный участок хромосомы донора окажется в цитоплазме реципиента, зависит от расстояния, которое отделяет этот участок от *oriT*. У генов, далеких от *oriT*, мало шансов быть перенесенными. Чем ближе к последовательности фактора F находится ген (если мерить расстояние по тому флангу, в направлении которого идет перенос), тем чаще можно встретить этот ген в рекомбинантных хромосомах реципиентов.

Таким образом, на хромосоме донора задан градиент частоты переноса: гены *E. coli* фактически выстроены в очередь для отправки в другой штамм. «Очередь» опоясывает хромосому, начинаясь от края последовательности F, где находится *oriT*.

В принципе на основании распределения частот появления генетических маркеров в стандартных условиях конъюгации можно построить карту хромосомы *E. coli*. Такие карты, с линейкой из ста делений по одной минуте, существуют; они строятся со ссылкой на конкретный штамм Hfr. (Сто минут, напомним, требуется на перенос целой хромосомы.) Точность карты для разных групп генетических маркеров будет зависеть от положения, которое занимает фактор F в хромосоме Hfr. Гены с высокой частотой переноса будут картированы с большей точностью.

16.8 Плазмида Ti, населяющая бактериальные клетки, является возбудителем корончатого галла у растений

Ключевые понятия:

- Клетки растений, инфицированные бактерией *A. tumefaciens*, могут трансформироваться в опухолевые.
- Опухолеродная трансформация растительной клетки происходит из-за плазмиды, хранящейся в бактерии.
- Эта плазмида несет в себе гены синтеза и метаболизма опинов (производных аргинина); данные гены экспрессируются в опухолях, вызванных бактерией.

Перестановки и амплификация последовательностей ДНК происходят, как правило, внутри одного генома. Однако у высших покрытосеменных двудольных растений есть ряд заболеваний, обусловленных вторжением плазмидной ДНК. У большинства двудольных растений такое заболевание, **корончатый галл**, может быть индуцировано почвенной бактерией *Agrobacterium tumefaciens*. На РИС. 16.13 представлена фотография больного растения, в котором завелась бактерия-паразит. Ее присутствие обернулось генетическим изменением одной из клеток растения. Это изменение благоприятно для паразита, но вредоносно для растения.



РИС. 16.13 Агробактерия, несущая нопалиновую плазмиду Ti, индуцировала тератому — опухоль, в которой различимы дифференцированные структуры, напоминающие зачатки органов растения. Фотография любезно предоставлена из имущества Jeff Schell. Используется с разрешения Max Planck Institute for Plant Breeding Research, Cologne.

Агробактерии индуцируют образование опухоли, после чего опухолевые клетки растут самостоятельно (то есть для аномального роста постоянное присутствие бактерии им не требуется). Опухоль появляется после того, как растительная клетка претерпевает трансформацию — за счет экспрессии генетического материала, полученного от агробактерии. Трансформированная растительная клетка имеет статус опухолеродной. Такая клетка имеет нестандартную программу управления собственным ростом и дифференцировкой.

Агробактерия может вызвать трансформацию растительной клетки, потому что несет в себе **плазмиду Ti** (от *англ.* tumor inducing plasmid — плазмида, индуцирующая опухоль), которая поддерживается в агробактерии как отдельный независимый репликон. В плазмиде Ti есть кластеры генов, которые участвуют в разнообразных аспектах жизнедеятельности бактериальной и раститель-

ной клеток. Гены одного из кластеров обусловливают перепрограммирование клеточного цикла растительной клетки. Есть также кластер так называемых «опиновых генов», задействованных в синтезе и усвоении особых органических веществ — **опинов**, производных аминокислоты аргинина.

Плазмиды Ti (и вместе с ними агробактерий, в чьей цитоплазме они содержатся) по типу опинов подразделяют на четыре группы.

- **Нопалиновые** плазмиды несут в себе гены для синтеза нопалинов в опухолях и для усвоения этих нопалинов бактериями. Нопалиновые опухоли могут дифференцироваться в виде трансформированных побегов. Так как они имеют тенденцию к морфогенезу, характерному для эмбрионального развития, нопалиновые опухоли, по аналогии с тератомами млекопитающих, называют **тератомами**.
- **Октопиновые** плазмиды похожи на нопалиновые, только с них продуцируется другой опин — октопин. Почему-то, как правило, октопиновые опухоли не дифференцируются и не образуют побегов-тератом.
- **Агропиновые** плазмиды несут в себе гены, задействованные в метаболизме агропина. Опухоли не дифференцируются, плохо развиваются и быстро отмирают.
- Агробактерии с **плазмидами Ri** индуцируют различные опухоли. У одних растений это типичный корончатый галл, у других растений появляются «волосатые корни» (hairy root disease). Кроме генов для метаболизма агропина плазмиды Ri могут содержать сегменты нопалиновых и октопиновых плазмид.

Классы генов, содержащихся в плазмидах Ti, сведены в таблицу на РИС. 16.14. Те из них, что экспрессируются в агробактерии (кроме опиновых генов), ответственны за репликацию плазмиды (включая контроль числа копий с сопутствующим эффектом несовместимости). Также эти гены ответственны за перенос плазмиды между бактериями, резистентность к фаговым инфекциям, синтез некоторых соединений (в частности, токсичных для других почвенных бактерий).

Экспрессирующиеся в клетке растения (кроме опиновых генов) гены плазмиды Ti ответственны за интеграцию материала плазмиды Ti в геном растения, за трансформацию клетки, за индукцию образования трансформированных побегов и корней (если они образуются).

Насколько специфичны продукты опиновых генов, зависит от плазмиды. Гены, которые необходимы для синтеза растительной клеткой конкретного опина, непременно сцеплены с генами, чьи продукты помогают агробактерии усваивать именно этот опин. Таким образом, любой штамм агробактерий имеет свой характерный набор Ti-маркеров и «располагает» опухолью, снабжающей его полезными опинами. Эти опины могут быть единственным источником углерода и/или азота, необходимым штамму-индуктору для поддержания жизнедеятельности. Трансформированная растительная клетка синтезирует те опины, которые может использовать бактерия; этот принцип неизменно соблюдается.

Гены Ti экспрессируются в бактериях и растениях

Локус	Функция	Ti-плазмида
vir	Перенос ДНК в растение	Все плазмиды Ti
shi	Индукция образования побегов	Все плазмиды Ti
roi	Индукция образования корней	Все плазмиды Ti
nos	Синтез нопалина	Нопалин
noc	Усвоение нопалина	Нопалин
ocs	Синтез октопина	Октопин
occ	Усвоение октопина	Октопин
tra	Перенос у бактерий	Все плазмиды Ti
Inc	«Гены несовместимости»	Все плазмиды Ti
oriV	Ориджин репликации	Все плазмиды Ti

РИС. 16.14 Гены плазмиды Ti участвуют в жизни бактериальной и растительной клеток

16.9 Плазмида Ti содержит гены, необходимые для вторжения Т-ДНК в геном растения

Ключевые понятия:

- Копия обширной области плазмиды Ti переносится в ядро растительной клетки.
- Гены *vir* плазмиды Ti нужны для осуществления переноса, однако находятся вне области, переносимой в растительную клетку.
- Экспрессия генов *vir* индуцибельна; она начинается под действием фенольных соединений, которые выделяет растение с поврежденной тканью.
- При связывании с веществом-индуктором мембранный белок VirA фосфорилирует сам себя по гистидину.
- VirA переносит фосфат на VirG, тем самым активируя его.
- Связка VirA-VirG представляет собой «фосфогистидиновое реле», типичный для бактерий компонент сигнального пути.

Взаимодействие агробактерии с растительной клеткой изображено на РИС. 16.15. Агробактерия, располагаясь вплотную к растительной клетке, передает ее ядру ДНК плазмиды Ti.

Область генома Ti, передаваемая растительной клетке, называют Т-областью. Копия Т-области, **Т-ДНК**, встраивается в геном растения и берет под свой контроль функцию синтеза опинов и трансформацию растительной клетки.

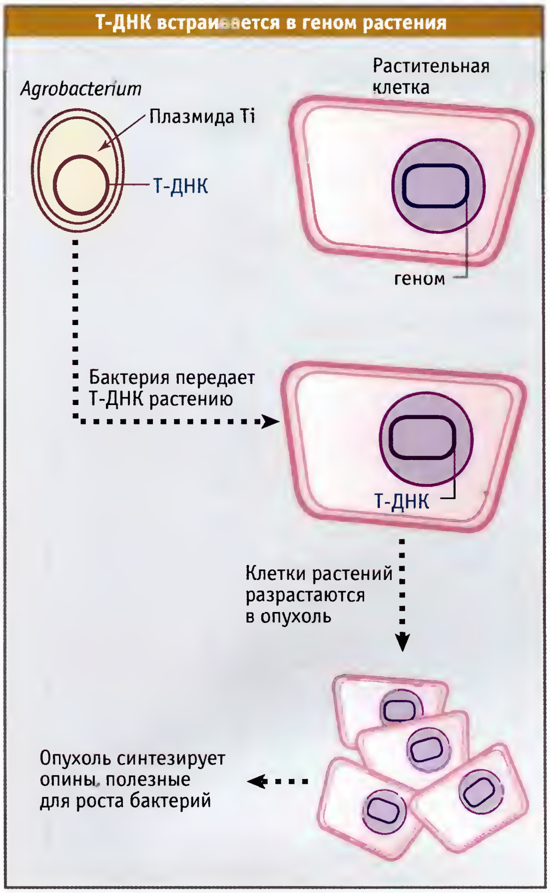


РИС. 16.15 Из агробактерии, несущей плазмиду Ti, в ядро растительной клетки переносится Т-ДНК. Из-за того, что Т-ДНК встраивается в геном клетки-хозяина и экспрессируется, растительная клетка претерпевает трансформацию

Трансформация растительной клетки в опухолеродную обеспечивается тремя группами локусов:

- Для начальных этапов внедрения бактерии в растительную клетку необходимы бактериальные белки (отвечают за синтез полисахаридов клеточной стенки бактерий). Они кодируются в локусах *chvA*, *chvB* и *pscA* на хромосоме агробактерии.
- Для высвобождения и инициации переноса Т-ДНК требуется область *vir* плазмиды Ti. Область *vir* находится за пределами Т-области.
- Для того чтобы растительная клетка претерпела трансформацию, нужна Т-ДНК.

С геномной организацией двух самых распространенных плазмид Ti можно ознакомиться по РИС. 16.16. Геномы нопалиновой и октопиновой плазмид Ti, по 200 тпн каждый, совпадают примерно на 30%. Общие участки содержат гены, участвующие в различных этапах взаимодействия агробактерии с растением-хозяином. Правда, расположение этих генов у нопалиновой и октопиновой плазмид довольно сильно отличается.

В каждой плазмиде примерно 23 тпн (из которых совпадают ~9 тпн) занимает Т-область. Внутри Т-области находятся гены синтеза опинов (*Nos* или *Ocs*); гены метаболизма соответствующих опинов (*Noc* или *Occ*) находятся за пределами Т-области. Нопалиновая и октопиновая плазмиды несут информацию о похожих, но не идентичных, морфогенетических функциях; для них характерны разные типы опухолей.

Таким образом, онкогенность агробактерий по отношению к растению-хозяину базируется не только в Т-области плазмиды Ti. В поддержании опухолеродного состояния трансформированной клетки, действительно, заняты только гены Т-области. Другие гены, расположенные за пределами этой области, могут иметь отношение к обретению клеткой статуса опухолеродной, однако для удержания клеткой этого статуса продукты таких генов не нужны. Они играют вспомогательную роль в переносе Т-ДНК в ядро растительной клетки, и, возможно, оптимизируют условия развития инфекции (например, баланс гормонов в инфицированной ткани). Некоторые мутации в геноме Ti специфичны относительно вида растения-хозяина, они приводят к образованию опухоли только в определенных видах растений.

Гены вирулентности *vir*, расположенные в Ti за пределами Т-области, кодируют белки, задействованные в процессе переноса Т-ДНК. Шесть локусов (*virA*, *-B*, *-C*, *-D*, *-E*, *-G*), разбросанные на интервале 40 тпн, имеют базовые характеристики, приведенные на РИС. 16.17. Каждый локус транскрибируется как отдельная единица; некоторые из локусов *vir* содержат более одной открытой рамки считывания.

Вторжение Т-ДНК в геном растения, если рассматривать этот процесс во всей полноте, можно разделить как минимум на два этапа:

- Агробактерия контактирует с растительной клеткой, индуцируя экспрессию генов *vir*.
- Продукты генов *vir* заставляют Т-ДНК перейти в ядро растительной клетки, где Т-ДНК встраивается в одну из хромосом.

По отношению к одному из этапов вторжения гены *vir* подразделяют на две группы. Так, *virA* и *virG* служат регуляторами; в ответ на изменение среды

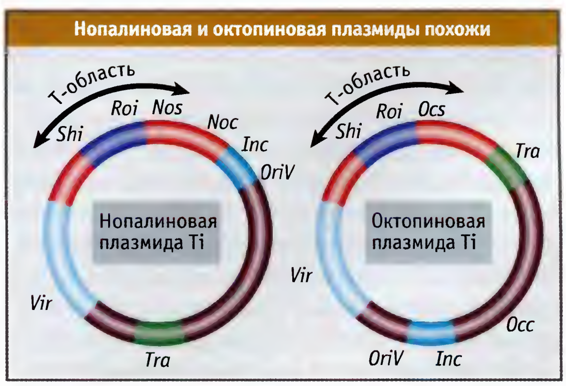


РИС. 16.16 Нопалиновая и октопиновая плазмиды Ti содержат разнообразные гены, в том числе гены Т-области со сходными функциями

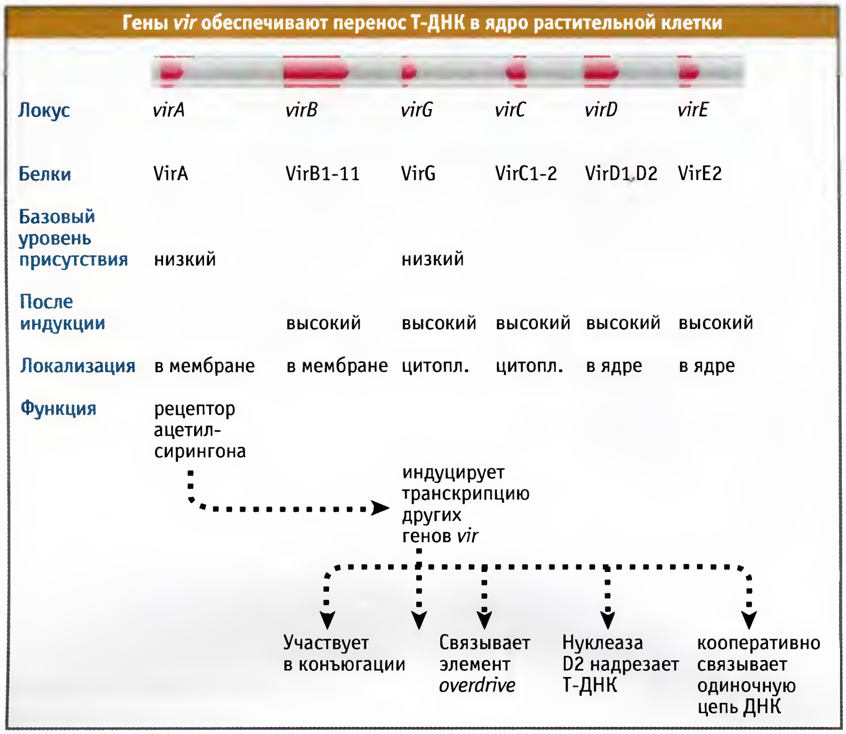


РИС. 16.17 **Область *vir* плазмиды Ti содержит шесть локусов, ответственных за перенос ДНК в инфицированное растение**

(ткани растения), благоприятствующее вторжению, продукты этих генов индуцируют экспрессию других генов *vir*. Мутанты с нарушениями функции *virA* и *virG* авирулентны, поскольку не способны экспрессировать остальные гены *vir* (*virB*, *virC*, *virD* и *virE*), кодирующие белки, непосредственно участвующие в переносе ДНК. Мутанты по генам *virB* и *virG* авирулентны у всех растений, тогда как эффекты мутаций, инактивирующих *virC* или *virE*, варьируют у разных видов растений-хозяев.

Гены *virA* и *virG* экспрессируются конститутивно (на довольно низком уровне). Их белки реагируют на сигнал в виде молекул, производных фенола, вырабатываемых растением в ответ на механическое повреждение. Формула одного из таких соединений изображена на РИС. 16.18; у табака (*Nicotiana tabacum*) их вырабатывается два — ацетосирингон и α-гидроксиацетосирингон. Они активируют белок VirA, тот передает сигнал белку VirG, который, в свою очередь, индуцирует (с нулевого уровня) экспрессию генов *virB*, *-C*, *-D*, *-E*. Биологическая природа индукции каскада *vir* дает ответ на вопрос, почему корончатый галл, как правило, поражает только поврежденные растения.

Белки VirA и VirG вместе образуют «белковую машину» классического для бактерий типа. В ответ на стимуляцию белок-сенсор, VirA, фосфорилирует сам себя (автофосфорилирование), после чего переносит фосфатную группу на белок-эффектор, VirG. Двухкомпонентная связка изображена на РИС. 16.19.



РИС. 16.18 **Ацетосирингон (4-ацетил-2,6-диметоксифенол) вырабатывается в поврежденной ткани *N. tabacum*. Это вещество индуцирует перенос Т-ДНК от агробактерий**

Аналогичная связка, в исполнении белков EnvZ и OmpR, участвует в ответе бактериальной клетки на осмотический шок. У генов *virA* и *envZ* есть гомологичные участки; гены *virG* и *ompR* вообще очень похожи, так что их белки явно работают по одной и той же схеме.

Белок VirA образует гомодимер; он находится во внутренней мембране оболочки бактерии. Когда в периплазме бактерии, между наружной и внутренней мембранами ее оболочки, появляются молекулы фенольных соединений-индукторов, VirA фосфори-

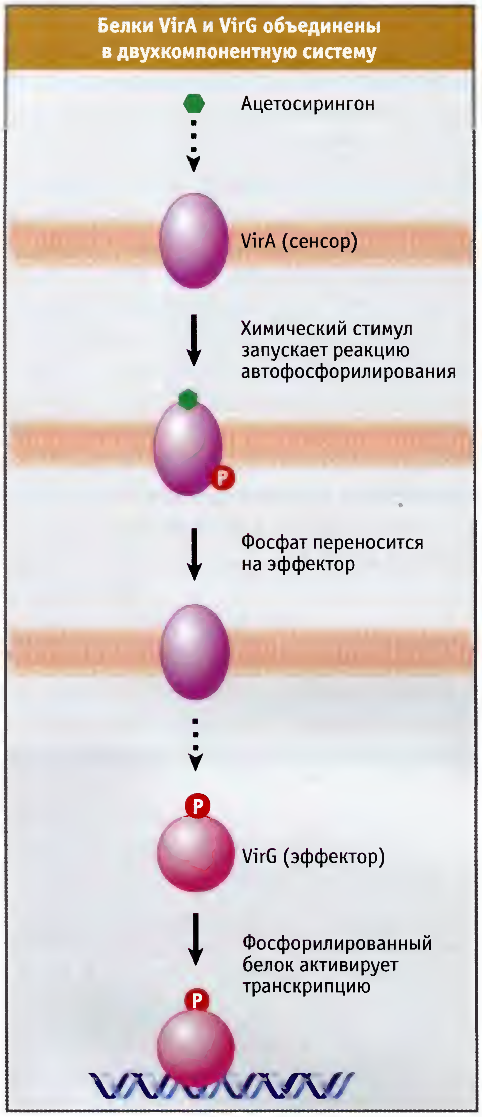


РИС. 16.19 **Двухкомпонентная система VirA-VirG отвечает на сигнал (появление фенольных соединений) активацией транскрипции генов-мишеней**

лирует сам себя по гистидину. После этого фосфатная группа переносится с гистидина (His) VirA на аспартат (Asp) VirG. Фосфорилированный VirG связывается с промоторами *virB*, *-C*, *-D*, *-E* и активирует транскрипцию. Кроме того, активация VirG приводит к транскрипции его гена, *virG*, с альтернативной стартовой точки (не с той, которая использовалась для конститутивной экспрессии). В результате количество белка VirG увеличивается, и передаваемый им сигнал усиливается.

Из остальных локусов *vir* лучше других изучен *virD*. Он содержит четыре открытых рамки считывания. Два продукта *virD* работают как эндонуклеаза, которая надрезает ДНК; ее мишень — конкретный сайт Т-области плазмиды Ti. Так инициируется процесс переноса Т-ДНК в растительную клетку.

16.10 Перенос Т-ДНК схож с конъюгацией бактерий

Ключевые понятия:

- Т-ДНК образуется путем подрезания дуплекса ДНК у правой границы Т-области.
- Во вскрытом дуплексе есть затравочный 3′-конец, с него начинается синтез новой цепи ДНК, сопровождающийся вытеснением одной из цепей дуплекса; вытесненный фрагмент одиночной цепи ДНК уходит из бактерии в ядро растительной клетки.
- Когда синтез новой цепи ДНК доходит до левой границы Т-области, перенос терминируется.
- Одноцепочечная Т-ДНК переносится в растительную клетку в комплексе с белком VirE2.
- После переноса Т-ДНК переходит в двухцепочечную форму и подлежит интеграции в геном растения.
- Несмотря на то что механизм интеграции изучен не полностью, Т-ДНК активно используется в лабораторной практике для доставки новых последовательностей в геном растительной клетки.

Процесс переноса генетического материала Ti в растительную клетку однонаправлен и строго избирателен; его единственным объектом является Т-ДНК. На РИС. 16.20 изображена Т-область нопалиновой плазмиды и превращения, которые с ней происходят. В интактной плазмиде Ti эта область окаймлена повторами по 25 пн (правый повтор отличается от левого двумя нуклеотидами). У копии Т-ДНК, интегрированной в геном растения, отсутствует почти весь правый повтор (от него остается 1 или 2 пн). Левая граница вставки задана не так четко: она может пройти и по левому повтору, и глубже него — по одному из нескольких возможных сайтов, разбросанных на участке ~100 пн. Иногда в геноме растения можно встретить множество тандемных копий Т-ДНК.

Модель переноса Т-ДНК в растительную клетку представлена на РИС. 16.21. В результате надреза правого повтора (то есть разреза одной из цепей ДНК) образуется свободный 3′-конец — затравка для синтеза новой цепи ДНК. Новая цепь ДНК вытесняет старую, которая переносится в ядро растительной клетки. Перенос терминируется после того, как синтез новой цепи ДНК достигнет надреза в левом повторе (сделанного по той же цепи ДНК, что и первый надрез). Эта модель хорошо объясняет особую значимость правого повтора в свете очевидной полярности всего процесса. Если левый повтор по каким-то причинам остается ненадрезанным, перенос может пойти дальше, с захватом материала левого фланга Т-области Ti.

Таким образом, перенос генетического материала в ядро растительной клетки обусловлен производством одной одноцепочечной молекулы ДНК в цитоплазме бактерии, которая заражает растение. Во время переноса эта молекула ДНК находится в комплексе с белком; все вместе иногда называют Т-комплексом. Белок, покрывающий Т-ДНК во время переноса, называется VirE2. Этот белок осуществляет кооперативное связывание одиночной цепи ДНК; в нем есть сигнал ядерной локализации.

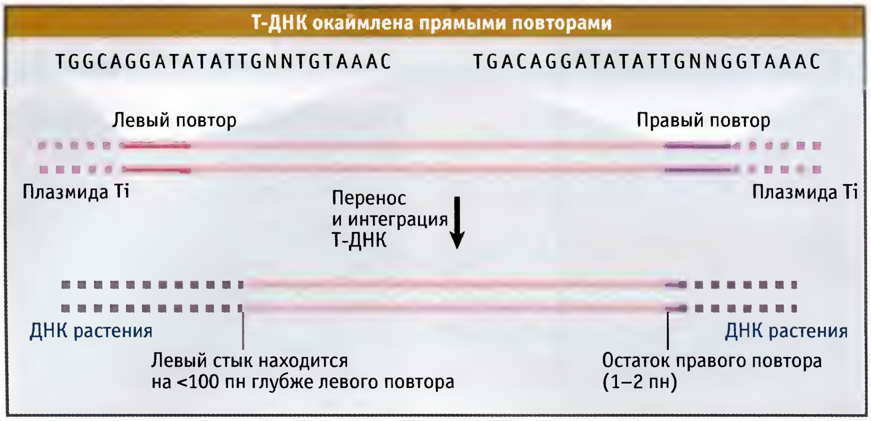


РИС. 16.20 Т-область плазмиды Ti окаймлена почти идентичными повторами по 25 пн. Правый повтор необходим для переноса Т-ДНК и ее интеграции в геном растения. Вставка Т-ДНК имеет фиксированную правую границу по началу правого повтора (от него остается 1 или 2 пн). Левая граница вставки не фиксирована и может быть заглублена почти на 100 пн

Белок VirE2 обеспечивает транспорт Т-ДНК в ядро растительной клетки. Во время переноса с 5′-концом Т-ДНК постоянно связан белок VirD2 (одна из субъединиц эндонуклеазы VirD1,2, вносящей инициирующий надрез в правый повтор Т-области Ti). В переносе участвует еще одиннадцать белков — продуктов оперона *virB*.

В плазмиде Ti (вне Т-области, вплотную к ее правой границе) находится короткая последовательность, которую называют *overdrive*. Она работает как очень мощный стимулятор переноса. По сути, *overdrive* — это *цис*-действующий энхансер; чтобы проявить себя, он должен находиться в одной молекуле ДНК с последовательностью-мишенью. Последовательность *overdrive* повышает эффективность переноса, даже если она локализована на расстоянии нескольких тпн от границы Т-области. С *overdrive* ассоциированы белки VirC1 и (возможно) VirC2.

Паттерн интеграции Т-ДНК в геном растения у разных плазмид разный. У октопиновых плазмид он сложнее, чем у нопалиновых. Сами вставки у октопиновых плазмид вариабельнее; для каждой октопиновой плазмиды можно выделить несколько разновидностей вставки, отличающихся целыми элементами последовательностей. Это говорит о том, что у Т-области октопиновой плазмиды есть несколько регуляторных элементов, которые могут служить мишенями для внесения надрезов и/или терминации синтеза ДНК.

На молекулярном уровне перенос Т-ДНК напоминает конъюгацию бактерий. Как мы помним, при конъюгации одиночная цепь ДНК (например, фактора F) переходит из одной клетки в другую. Неудивительно, что гены оперона *virB* гомологичны генам *tra* фактора F и некоторых других плазмид-операторов конъюгации (см. разд. 16.7 «При конъюгации происходит перенос одноцепочечной ДНК»). Между переносом ДНК при

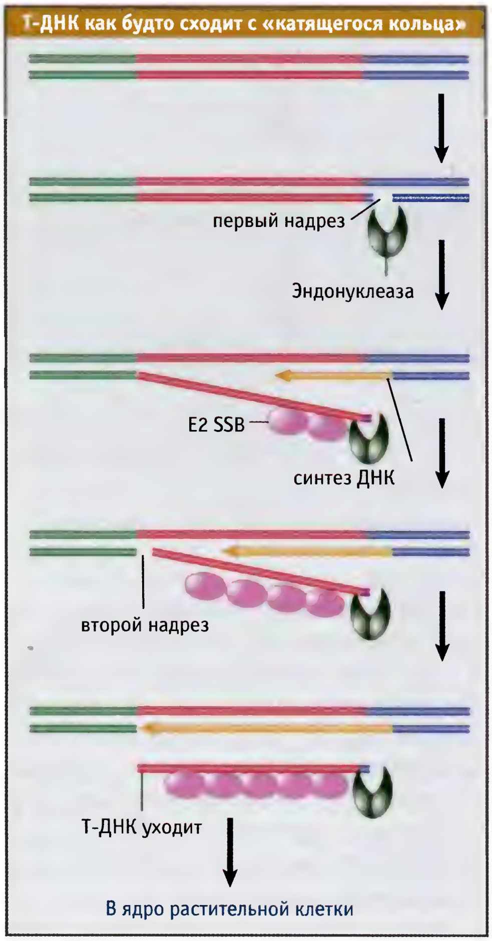


РИС. 16.21 Т-ДНК образуется в результате смещения одной из цепей дуплекса за счет синтеза новой цепи ДНК. Синтез начинается на месте надреза в правом повторе. Реакция терминируется надрезом по левому повтору

конъюгации и переносом Т-ДНК можно усмотреть следующую разницу: Т-ДНК ограничена левой границей Т-области Ti, а перенос ДНК при конъюгации бактерий, с точки зрения нуклеотидной последовательности, ничем не ограничивается.

Каким именно способом Т-ДНК внедряется в геном растения, мы не знаем. Ясно, что в какой-то момент перед интеграцией одиночная цепь Т-ДНК должна быть переведена в форму дуплекса. В инфицированных растительных клетках можно обнаружить кольцевые Т-ДНК; полагают, что они являются продуктами рекомбинации между левым и правым 25 пн повторами. Может быть, эти кольцевые Т-ДНК являются промежуточными продуктами очень конкретной реакции, о которой, тем не менее, пока ничего не известно. Само событие интеграции Т-ДНК в геном растения представляет собой какую-то форму негомологичной рекомбинации, поскольку между Т-ДНК и сайтами интеграции нет никакой выраженной гомологии.

Насколько полноценна внедрившаяся Т-ДНК как генетическая единица? Сколько копий Т-ДНК внедряется в один геном? Какие сайты геномной ДНК растения подходят для внедрения Т-ДНК? Подлежит ли экспрессия Т-генов регуляции со стороны элементов генома растения, или же эти гены подчиняются только своим собственным системам контроля? Эти вопросы принципиальны для понимания сущности трансформации растительной клетки под влиянием генетического вмешательства со стороны плазмиды Ti.

Какими отличительными особенностями обладает сайт-мишень в геноме растения? Геномные последовательности, фланкирующие вставку Т-ДНК, содержат много пар А-Т (что характерно для сайтов-мишеней некоторых транспозонов). Структуру каждого отдельно взятого сайта мишени довольно трудно анализировать из-за того, что при внедрении Т-ДНК в окрестностях ее концов происходят перестановки сегментов нуклеотидной последовательности. Неясно также, возникают ли при внедрении Т-ДНК новые повторы подобно тому, как это происходит при транспозиции мобильных элементов генома.

Как мы знаем, ни один из генов Т-области не задействован в переносе Т-ДНК. После того как перенос закончился и внедрение состоялось, гены Т-области (теперь уже Т-области генома растения) начинают экспрессироваться. Область содержит несколько транскрипционных единиц; каждая из них несет ген со своим собственным промотором. Функции генов Т-области влияют на состояние растительной клетки; они направлены на поддержание опухолеродного потенциала клетки, контроль морфогенеза дополнительных побегов и корней, подавление выхода из клеточного цикла и терминальной дифференцировки.

В заключение еще раз заострим внимание на организации плазмиды Ti. За пределами Т-области находятся гены, необходимые для инициации у растения онкогенеза. Большинство этих генов задействовано в переносе Т-ДНК. Кроме «генов переноса» в плазмиде Ti (также за пределами Т-области) есть гены, позволяющие агробактерии усваивать опин, продуцируемый растительной клеткой после трансформации. Наконец, внутри Т-области находятся гены контроля роста опухоли, а также гены синтеза опинов для поддержания агробактерии, поставляющей Т-ДНК.

С точки зрения практики, способность агробактерий к генетической трансформации растительной клетки равносильна возможности вносить в растения новые гены. Как мы видим, в плазмиде Ti функции переноса/интеграции и онкогенеза заданы раздельно. Эта особенность позволяет конструировать новые Ti-плазмиды, онкогенетические функции которых заменены другими генами (например, для того чтобы проверить, какой эффект окажет появление у растения чужеродных генов). Можно сказать, что в сфере генной инженерии растений замечательная естественная система доставки генов пришлась как нельзя кстати.

16.11 Резюме

Образование структуры под названием «катящееся кольцо» характерно для альтернативной формы репликации кольцевых молекул ДНК. Репликацию по типу «катящегося кольца» инициирует внесение разреза в одну из цепей ДНК двухцепочечного ориджина репликации. Возникающий при этом свободный 3′-конец используется как затравка для синтеза новой цепи ДНК. Эта новая цепь ДНК смещает с матрицы ее исходного партнера; смещенная цепь ДНК представляет собой продукт репликации. Каждый оборот «катящегося кольца» эквивалентен продукции новой одноцепочечной линейной копии генома.

Репликация по типу «катящегося кольца» свойственна некоторым фагам. Один из них, фаг φX174, имеет специальный белок А, который разрезает одну из цепей ориджина. *In vivo* белок А проявляет необычную *цис*-действующую активность, взаимодействуя только с той ДНК, которая послужила матрицей для его синтеза. Белок А остается прикрепленным к смещенной цепи ДНК — до тех пор, пока репликация не пройдет полный круг (то есть пока не будет синтезирована полноразмерная копия фагового генома). Затем белок А снова разрезает ту же цепь по такому же сайту в ориджине. При этом полноразмерная одноцепочечная копия генома превращается в кольцевую молекулу и уходит, а репликация начинает новый круг (иными словами, стартует новый раунд репликации).

«Катящиеся кольца» можно также наблюдать при конъюгации бактерий. Иногда две бактериальные клетки тесно взаимодействуют, чтобы одна из них, донор, могла передать другой, реципиенту, свою плазмиду (например, фактор F). Контакт между донором и реципиентом инициируется при помощи специальных ворсинок на поверхности клетки-донора. Бывает, что геном фактора F

находится в составе бактериальной хромосомы. Штамм бактерий с такой хромосомой называют Hfr. Он может делиться со штаммом-реципиентом *собственным* генетическим материалом. При конъюгации донор использует репликацию ДНК, чтобы восполнить утрату одиночной цепи ДНК, покидающей клетку. Эта одиночная цепь попадает в цитоплазму реципиента и там, опять же при помощи репликации, переводится в двухцепочечную форму. Репликация *не* является движущей силой переноса ДНК при конъюгации.

Бактерии *Agrobacterium tumefaciens*, агробактерии, индуцируют образование опухоли в поврежденной ткани растения. Если ткань растения механически повреждена, ее клетки секретируют фенольные соединения, активирующие гены *vir* в плазмиде Ti, которая находится в агробактерии. Продукты экспрессии генов *vir* заставляют одну из цепей ДНК Т-области плазмиды Ti, Т-ДНК, перейти в ядро растительной клетки. Перенос инициируется на одной из границ Т-области и терминируется по одному из сайтов возле другой границы. Одноцепочечная Т-ДНК переводится в двухцепочечную форму и внедряется в геном растения. Экспрессия генов, находящихся в Т-ДНК, приводит к опухолеродной трансформации растительной клетки, а также заставляет эту клетку продуцировать определенные разновидности опинов (органических соединений, производных аргинина). Другие гены плазмиды Ti (не попадающие в растительную клетку) помогают аггробактериям усваивать именно эти опины. На основе Т-ДНК были созданы векторы для внесения в геномы растений новых нуклеотидных последовательностей.

Литература

16.4 С «катящихся колец» сходят мультимеры репликонов

Исследование

Gilbert, W. and Dressler, D. (1968). DNA replication: the rolling circle model. *Cold Spring Harbor Symp. Quant. Biol.* 33, 473–484.

16.6 Фактор F передается от бактерии к бактерии при конъюгации

Исследование

Ihler, G. and Rupp, W. D. (1969). Strand-specific transfer of donor DNA during conjugation in *E. coli. Proc. Natl. Acad. Sci. USA* 63, 138–143.

16.7 При конъюгации происходит перенос одноцепочечной ДНК

Обзор

Frost, L. S., Ippen-Ihler, K., and Skurray, R. A. (1994). Analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiol. Rev.* 58, 162–210.

Ippen-Ihler, K. A. and Minkley, E. G. (1986). The conjugation system of F, the fertility factor of *E. coli. Annu. Rev. Genet.* 20, 593–624.

Lanka, E. and Wilkins, B. M. (1995). DNA processing reactions in bacterial conjugation. *Annu. Rev. Biochem.* 64, 141–169.

Willetts, N. and Skurray, R. (1987). Structure and function of the F factor and mechanism of conjugation. In Neidhardt, F. C., ed. *Escherichia coli* and *Salmonella typhimurium.* Washington, DC: American Society for Microbiology, pp. 1110–1133.

Взаимосвязь между репликацией и клеточным циклом у бактерий

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

- У плазмид есть системы сайт-специфичной рекомбинации, необходимые для восстановления мономеров за счет внутримолекулярной рекомбинации.
- У плазмид также есть системы сегрегации, которые гарантируют распределение дуплицированных плазмид по разным дочерним клеткам.

17.10 Несовместимость плазмид обусловлена сходством их репликонов

- Плазмиды, принадлежащие к одной группе совместимости, имеют ориджины, регулируемые одной и той же системой контроля.

17.11 Система совместимости ColE1 контролируется РНК-регулятором

- Для репликации плазмиды ColE1 необходимо, чтобы сначала через ее ориджин прошла транскрипция; продукт транскрипции расщепляется РНКазой H с образованием затравочного конца.
- Регуляторная РНК I представляет собой короткую антисмысловую РНК; она взаимодействует с транскриптом и препятствует расщеплению, порождающему затравочный конец.
- Белок Rom усиливает взаимодействие между РНК I и транскриптом.

17.12 Как происходит репликация и сегрегация ДНК в митохондриях эукариот?

- Репликация и распределение мтДНК по дочерним митохондриям носят стохастический характер.
- Сами митохондрии распределяются по дочерним клеткам тоже произвольным образом.

17.13 Резюме

17.1 Введение

Прокариоты и эукариоты очень сильно отличаются друг от друга по способу контроля соответствия между событием репликации ДНК и развитием клеточного цикла. Для всех эукариот характерно, что:

- хромосомы находятся в ядре,
- каждая хромосома состоит из многих репликонов,
- репликация представляет собой координированное воспроизводство ДНК всех репликонов в продолжение дискретного периода клеточного цикла,
- решение о начале репликации выносится эффекторами сложного сигнального пути, регулирующего клеточный цикл.

Кроме того,

- во время митоза дуплицированные хромосомы распределяются (сегрегируют) по дочерним клеткам благодаря действию специального аппарата.

У бактерий запуск репликации контролируется по-другому. Когда масса бактерии в своем увеличении переходит пороговый уровень, в единственной области ориджина бактериальной хромосомы инициируется запуск репликации; это иллюстрирует РИС. 17.1. Сегрегация дочерних хромосом бактерии направлена на то, чтобы они оказались по разные стороны перегородки-септы, образующейся между дочерними клетками при делении.

Каким образом бактериальная клетка узнает, что пора начать раунд репликации? Событие инициации происходит каждый раз при одних и тех же значениях отношения клеточной массы к числу хромосомных ориджинов. Клетки, которые растут быстрее, достигают более крупных размеров и содержат большее число ориджинов. Рост *E. coli* можно измерить в **единичных клетках** (unit cell). Под единичной клеткой подразумевается целостный клеточный сегмент длиной 1,7 мкм. Бактерия содержит один ориджин репликации на единичную клетку; быстро растущая клетка с двумя ориджинами имеет длину от 1,7 до 3,4 мкм.

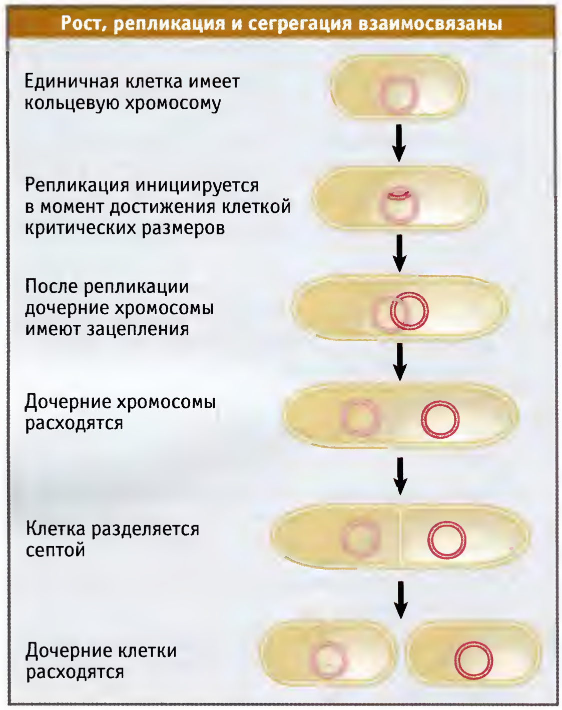


РИС. 17.1 Когда размер бактериальной клетки достигает критического порогового значения, в ориджине ее хромосомы инициируется репликация. Дочерние хромосомы, порожденные репликацией, могут оказаться сцепленными из-за успевшей произойти рекомбинации или могут иметь зацепления (как конкатамеры). Дочерние хромосомы должны быть освобождены друг от друга; после этого они расходятся и во время деления находятся по разные стороны перегородки-септы

Каким образом клетка контролирует свою массу? Можно предположить, что существует гипотетический белок-инициатор, непрерывно синтезируемый на всем протяжении клеточного цикла; накопление этого белка в критическом количестве срабатывает на запуск инициации репликации. В эту гипотезу хо-

рошо вписывается необходимость синтеза белка для события инициации. Другое возможное объяснение осведомленности клетки о своей массе состоит в том, что в некоторый фиксированный момент клеточного цикла (например, при делении) в клетке происходит синтез белка-ингибитора, который, по мере роста (увеличения объема) клетки, разбавляется до концентрации ниже эффективной.

Разделение дочерних хромосом бактерии после репликации — отдельная задача; некоторые сложности в разделении дочерних хромосом возникают из-за их кольцевой природы. О кольцевых хромосомах говорят, что они *катенированы*, если их кольца проходят друг сквозь друга, на манер цепных звеньев, и хромосомы оказываются соединенными. Чтобы их разъединить, нужны топоизомеразы. Альтернативный тип структуры формируется за счет события рекомбинации: акт рекомбинации между двумя мономерами преобразует их в единый димер. Этот процесс осуществляет специализированная система рекомбинации, воссоздающая независимые мономеры. Процессом разделения конкатамеров, по сути, заведуют ферментные системы, нацеленные непосредственно на работу с конкретными последовательностями ДНК.

17.2 Репликация взаимосвязана с клеточным циклом

Ключевые понятия:

- Время удвоения числа бактерий в культуре *E. coli* может варьировать в более чем десятикратном диапазоне, в зависимости от условий роста.
- Чтобы реплицировать бактериальную хромосому (при нормальной температуре), требуется 40 мин.
- Завершение раунда репликации срабатывает на запуск деления бактериальной клетки 20 мин спустя.
- Если время удвоения составляет менее 60 мин, следующий раунд репликации должен начаться раньше, чем произойдет деление, запущенное предыдущим раундом.
- Поэтому быстрые темпы роста сопровождаются у бактерий появлением «хромосом с множеством вилок».
- Раунд репликации инициируется по достижении определенного значения отношения клеточной массы к числу хромосомных ориджинов.
- На единичную клетку длиной 1,7 мкм приходится один ориджин.

У бактерий между репликацией и клеточным циклом есть два элемента обратной связи:

- частота инициации раундов репликации подстраивается под скорость роста клетки,
- завершение раунда репликации срабатывает (с задержкой) на запуск деления клетки.

Скорость роста бактерий в культуре характеризуется **временем удвоения**, то есть продолжительностью периода, необходимого клеткам, чтобы удвоить свою численность. Чем короче время удвоения, тем быстрее растут клетки. Скорость роста клеток *E. coli* характеризуется временами удвоения в диапазоне от 18 мин до 180 мин. Бактериальная хромосома представляет собой одиночный репликон; таким образом, частота раундов ее репликации определяется частотой событий инициации в единственном ориджине. Раунд репликации можно описать двумя константами:

- C, фиксированное время, которое требуется, чтобы реплицировать хромосому целиком. Составляет около 40 мин, что соответствует продвижению репликативной вилки со скоростью ~50 000 пн в мину. (При постоянной температуре ДНК синтезируется с более или менее стабильной скоростью, если только синтез не лимитируется дефицитом мономеров-предшественников.)
- D, фиксированное время, которое должно пройти от завершения репликации до клеточного деления, спровоцированного окончанием репликации. D составляет ~20 мин. Скорее всего, этот период времени нужен для сборки компонентов, необходимых для деления.

Константы C и D можно рассматривать как пределы возможностей клетки по части быстроты завершения соответствующих процессов; они соблюдаются при условии, что оба процесса — синтез ДНК и подготовка к делению — идут с максимально возможной скоростью. Указанные значения C и D применимы ко всем временам удвоения в диапазоне от 18 до 60 мин, но если клеточный цикл (промежуток между делениями) занимает более 60 мин, то обе фазы становятся более продолжительными.

Раунд репликации хромосомы должен быть инициирован задолго до того, как клетка разделится; минимальное время, которое должно пройти между этими двумя событиями, составляет: C + D = 60 мин. Для бактерий, которые делятся чаще, чем раз в 60 мин, раунд репликации должен быть инициирован прежде, чем закончится предыдущий клеточный цикл. Клетка, если можно так образно выразиться, рождается уже беременной следующим поколением.

Рассмотрим такой пример: клетки в культуре делятся каждые 35 мин. Тогда раунд репликации, приуроченный к конкретному делению, должен быть инициирован на 25 мин раньше, чем произойдет предыдущее деление. Эту ситуацию иллюстрирует РИС. 17.2, где комплектация бактериальной хромосомы отслежена на промежутке одного клеточного цикла (от деления до деления) как серия состояний, взятых с интервалом в 5 мин.

В момент деления (35/0 мин) каждая дочерняя клетка получает частично реплицированную хромосому. Репликативная вилка в такой хромосоме продолжает свое движение. На десятой мине, когда первая репликативная вилка еще не достигла участка терминации, в обоих ориджинах частично реплицированной хромосомы происходит инициация нового раунда репликации. Старт двух новых репликативных вилок превращает хромосому в **хромосому с множеством вилок**.



РИС. 17.2 У бактерий есть фиксированный временной интервал между инициацией репликации и клеточным делением; у *Escherichia coli* он составляет 60 мин. Поэтому, когда клетки быстро растут в культуре, их хромосомы «обрастают гроздьями» репликативных вилок. На схеме показаны репликативные вилки для одного из двух направлений репликации. На самом деле, как мы знаем, репликация кольцевой бактериальной хромосомы осуществляется двумя симметричными наборами репликативных вилок, движущихся в противоположных направлениях

В 15 мин — то есть за 20 мин до следующего деления — первая (полученная клеткой еще при делении) репликативная вилка достигает участка терминации. Конец ее существования знаменует окончание раунда репликации и позволяет дочерним хромосомам отделиться друг от друга (сепарировать); при этом каждая из них уже частично реплицирована новыми репликативными вилками (возникшими уже в этом клеточном цикле). Эти вилки продолжают свое движение по ДНК.

В момент деления две частично реплицированные хромосомы сегрегируют. При этом воссоздается ситуация, с которой мы начинали рассмотрение. Через 15 мин единственная репликативная вилка терминируется (успев за это время стать не единственной), а еще спустя 20 мин происходит деление. Мы видим, что событие инициации раунда репликации происходит за $1^{25}/_{35}$ клеточного цикла до события деления, с которым ассоциировано завершение этого раунда репликации.

Общий принцип зависимости между инициацией репликации и клеточным циклом состоит в том, что по мере ускорения роста клеток (и укорачивания клеточного цикла) событие инициации начинает опережать деление, которому оно предназначается, на все возрастающее число клеточных циклов. Содержание хромосомной ДНК в каждой отдельно взятой бактерии при этом увеличивается. Такую взаимосвязь можно расценивать как ответ клетки на свою неспособность сократить периоды C и D, чтобы привести темпы репликации в соответствие требованиям более короткого клеточного цикла.

17.3 Септа разделяет бактериальную клетку на две дочерние, каждая из которых содержит хромосому

Ключевые понятия:

- Образование септы инициируется по месту перисептального кольца, которое опоясывает клетку и представляет собой зону видоизмененной клеточной оболочки.
- Появление новых колец инициируется на серединах расстояний от септы до концов бактерии.
- Когда бактерия делится, в каждой дочерней клетке перисептальное кольцо находится посередине.
- Образование септы начинается по достижении клеткой определенной длины.
- Септа состоит из тех же пептидогликанов, что и оболочка бактерии.

Сегрегация хромосом у бактерий интересна тем, что механизм распределения хромосом по дочерним клеткам у них основан на участии самой ДНК. Это контрастирует с ситуацией в эукариотических клетках, где сегрегация достигается силами сложного митотического аппарата. Бактериальные хромосомы распределяются по дочерним клеткам с большой аккуратностью: в популяции (культуре) бактерий можно встретить безъядерные клетки, но они составляют менее 0,03% всех клеток.

Деление тела клетки у бактерий выполняется за счет образования перегородки, или **септы**. Септа образуется посередине клетки как впячивание (инвагинация) оболочки, которой окружена бактерия. Септа создает непроницаемый барьер между двумя частями клетки; она служит тем местом, по которому две дочерние клетки со временем полностью отделятся друг от друга. По поводу роли, которую септа играет в делении, можно задать два — на самом деле, родственных — вопроса: чем определяется местоположение септы, и чем гарантировано местонахождение дочерних хромосом по разные стороны от нее?

Образованию септы предшествует возникновение перисептального кольца. У *E. coli* или *Salmonella typhimurium* перисептальное кольцо можно наблюдать как зону, в которой оболочка бактерии видоизменена таким образом, что ее внутренняя мембрана теснее контактирует с клеточной стенкой и наружной мембраной. Эта зона опоясывает клетку (отчего и получила название «кольца»). Возникновение и дальнейшая судьба таких колец показаны на РИС. 17.3.

У каждой новой клетки перисептальное кольцо расположено по центру. С ростом клетки происходят следующие события: по каждую сторону от исходного кольца формируется новое; образовавшись вблизи исходного кольца, новые кольца сдвигаются в направлении концов бактерии на позиции, соответствующие ¼ и ¾ длины клетки; по центру клетки, как намечено исходным кольцом, образуется септа.

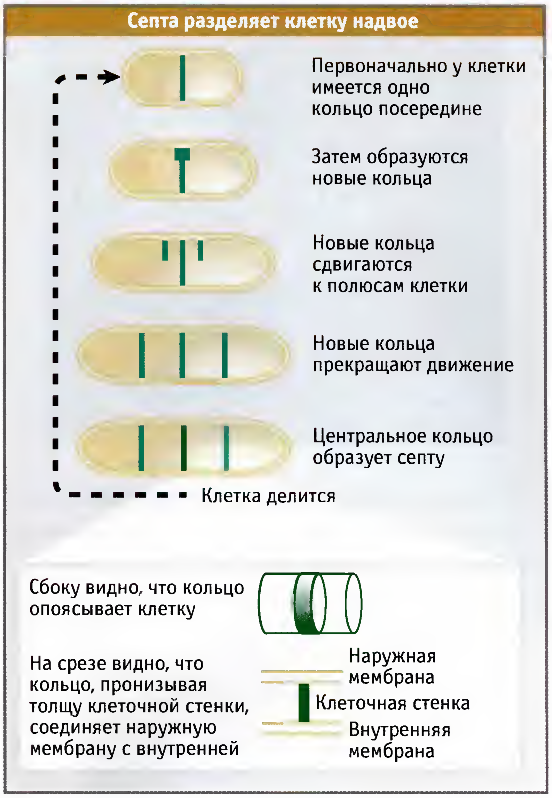


РИС. 17.3 **Дупликация и смещение перисептальных колец задают паттерн образования септ, делящих клетки**

После того как произойдет деление, расположение новых колец в дочерних клетках окажется центральным. Сдвиг перисептального кольца на правильную позицию критичен как гарантия деления клетки на одинаковые по размеру дочерние клетки. Механизм движения колец неизвестен.

Образование септы (септация) начинается по достижении клеткой фиксированной длины, 2L; при этом расстояние между новыми перисептальными кольцами в норме всегда равняется L. Мы не знаем, каким образом клетка определяет длину, но она, по всей видимости, измеряет ее как таковую (не площадь сечения и не объем).

Септа состоит из тех же компонентов, что и оболочка бактерии: жесткий слой пептидогликанов находится в периплазме, между внутренней и наружной мембранами. Пептидогликаны образуются в результате полимеризации три- или пентапептид-дисахаридных субъединиц, за счет реакций транспептидации и трансгликозилирования с использованием обеих биохимических составляющих такой субъединицы, олигопептидной и углеводной.

Палочковидная форма бактерии поддерживается энзиматическими активностями белков PBP2 и RodA. Они являются продуктами одного оперона, а также взаимодействуют друг с другом. Белок RodA принадлежит к семейству SEDS (название семейства происходит от слов shape, elongation, division и sporulation — форма, удлинение, деление и споруляция). Семейство SEDS представлено у всех бактерий, чьи клетки имеют пептидогликановую стенку. Каждый белок из семейства SEDS функционирует совместно со специфичной к нему транспептидазой, катализирующей образование сшивок в молекуле пептидогликана. Пенициллин-связывающий белок 2 (PBP2) как раз и является такой транспептидазой, взаимодействующей с белком RodA. Мутации в генах этих двух белков (любого из них в отдельности) приводят к потере бактериями своей вытянутой формы; они становятся круглыми. Здесь дает себя знать важный принцип, согласно которому форма и жесткость могут быть детерминированы за счет простого наращивания полимерной структуры.

За наработку пептидогликана септы отвечает другой фермент (см. разд. 17.5 «Для образования септы необходим белок FtsZ»). Септа исходно формируется как два слоя пептидогликанов; над расщеплением ковалентных сшивок между этими слоями работает белок EnvA; в результате дочерние клетки могут разойтись (сепарировать).

Возникновение и сдвиг перисептальных колец по ходу клеточного цикла говорят о том, что у бактерии механизм замера отрезков длины в пространстве ассоциирован с оболочкой. Логично предположить, что оболочка может также использоваться бактериальной клеткой как навигатор для сегрегации хромосом. Гарантированная сегрегация может обеспечиваться за счет непосредственной связи ДНК с клеточной мембраной. Если дочерние хромосомы прикреплены к мембране в соответствии с разметкой на будущие дочерние клетки, то образующейся септе ничего другого не остается, как разделить их физически. На Рис. [illegible] показано, как образование септы закрепляет распределение по разным дочерним клеткам хромосом, чьи ориджины привязаны к сайтам клеточной мембраны по разные стороны перисептального кольца.

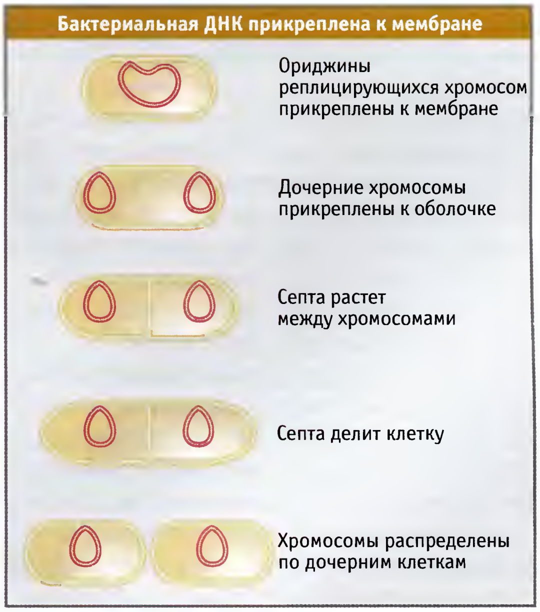


РИС. 17.4 **Механизм сегрегации может строиться на факте прикрепления бактериальной ДНК к мембране**

17.4 Мутации, нарушающие деление или сегрегацию, сказываются на форме клеток

Ключевые понятия:

- *fts*-мутанты формируют длинные нити, поскольку у них не образуются септы, разделяющие дочерние клетки.
- Мутанты, у которых септ слишком много, продуцируют мини-клетки; они мелкие и не содержат ДНК.
- Мутанты, у которых нарушена сегрегация дуплицированных хромосом, генерируют безъядерные клетки нормального размера.

Трудность получения мутантов, у которых нарушен процесс деления, проистекает от весьма вероятной летальности и/или плейотропности мутаций, затрагивающих столь актуальную функцию, как деление. Например, если возникновение перисептального кольца происходит в месте, которое вместе с тем очень существенно для роста оболочки бактерии, то отличить мутации, специфически препятствующие образованию кольца, от мутаций, замедляющих общие темпы роста оболочки, будет нелегко. Носители большинства мутаций, нарушающих работу аппарата деления, были идентифицированы как условные мутанты (conditional mutants), у которых процесс деления оказывается нарушен в результате наступления «неразрешающих», или непермиссивных, условий (nonpermissive conditions); чаще всего такие мутанты чувствительны к изменению температуры.

Мутации, нарушающие процесс деления или сегрегацию хромосом, приводят к сокрушительным изменениям в фенотипе. На рис. 17.5 и 17.6 показаны последствия, соответственно, неудавшегося процесса деления и неудачной сегрегации:

- Длинные нити, или филаменты, формируются, когда образование септы нарушено, а репликация хромосом идет нормальным образом. Бактерии продолжают расти и даже продолжают сегрегировать свои дочерние хромосомы, однако септ не образуется. Из-за этого клетки становятся очень длинными, нитевидными; по длине каждой такой клетки-филамента с регулярными интервалами распределены нуклеоиды — бактериальные хромосомы. Такой фенотип бывает у мутантов *fts*⁻, страдающих дефектом или множеством дефектов, в самом процессе деления (от *англ.* filamentation temperature sensitive — филаментация, чувствительная к температуре).
- Бактериальные **мини-клетки** (minicells) появляются, если образование септ происходит слишком часто или в неположенном месте; в результате одна из дочерних клеток не имеет хромосомы. Такая мини-клетка отличается небольшими размерами, и в ней нет ДНК; во всем остальном она выглядит нормальной.
- **Безъядерные** клетки (anucleate cells) образуются из-за аберрантной сегрегации; они, подобно мини-клеткам, не имеют хромосомы, но, поскольку септы в данном случае образуются как в норме, размер таких клеток стандартный. Порождение безъядерных клеток характерно для мутантов *par* (от partition, разделение), получивших свое название за дефекты сегрегации хромосом.

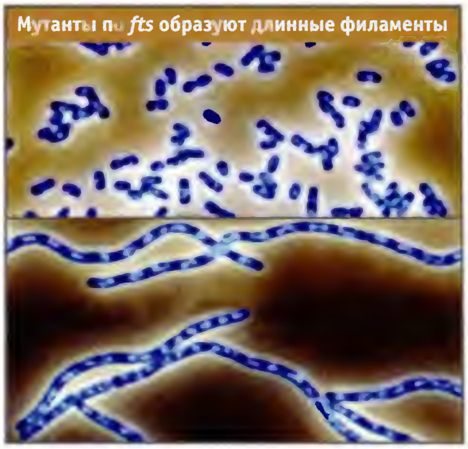


РИС. 17.5 *Вверху:* клетки дикого типа. *Внизу:* клетки-мутанты в непермиссивных температурных условиях не способны делиться. Это приводит к образованию многоядерных филаментов. Фотография любезно предоставлена Sota Hiraga, Kyoto University



РИС. 17.6 При нарушениях сегрегации хромосом в культуре *E. coli* начинают появляться безъядерные клетки. На фотографии клетки, где есть хромосома, окрашены в голубой цвет; дочерние клетки, не имеющие хромосомы, выглядят бесцветными. Фотография любезно предоставлена Sota Hiraga, Kyoto University

17.5 Для образования септы необходим белок FtsZ

Ключевые понятия:

- Продукт гена *ftsZ* требуется для образования септы в заранее намеченном месте.
- Белок FtsZ является ГТФазой, он образует кольцо на внутренней стороне оболочки бактерии и состоит в контакте с другими компонентами цитоскелета.

Ген *ftsZ* играет в делении одну из главных ролей. Мутации по *ftsZ* блокируют образование септ, что приводит к превращению бактерий в филаменты. Избыточная экспрессия *ftsZ* приводит к появлению мини-клеток, поскольку при этом на единичную клетку приходится слишком много событий септации. Мутации в *ftsZ* дают себя знать на разных стадиях, начиная от смещения перисептальных колец и заканчивая морфогенезом септы. Все наблюдения говорят о том, что белок FtsZ нужен, чтобы место, заранее намеченное перисептальным кольцом (periseptal annulus), было использовано для образования септы. При этом сам FtsZ не влияет на образование перисептальных колец, ни на их локализацию.

Белок FtsZ функционирует на ранней стадии образования септы. Перед началом деления белок FtsZ равномерно распределен в цитоплазме. По мере того, как клетка удлиняется и начинает перетягиваться посередине, область локализации FtsZ принимает вид кольца, опоясывающего клетку. Соответствующую структуру иногда называют **Z-кольцом**. На РИС. 17.7 можно разглядеть, что Z-кольцо занимает то же положение, что и центральное перисептальное кольцо с рис. 17.3. Образование Z-кольца — этап, лимитирующий темпы образования септы. В клеточном цикле Z-кольцо возникает по центру клетки вскоре после деления (в самом типичном случае через 1–5 мин), какое-то время (15 мин) держится, а затем быстро стягивается, пережимая клетку надвое.

По своей структуре белок FtsZ напоминает тубулин; это говорит о том, что образование Z-кольца может напоминать сборку микротрубочек в клетках эукариот. FtsZ обладает ГТФазной активностью; энергия от расщепления ГТФ расходуется на поддержку олигомеризации мономеров FtsZ для включения в состав кольца. Являясь динамической структурой, Z-кольцо непрерывно обменивается субъединицами с цитоплазматическим пулом FtsZ.

С FtsZ напрямую взаимодействуют два белка, ZipA и FtsA; они делают это независимо. Белок ZipA встроен во внутреннюю мембрану оболочки бактерии. Он служит средством прикрепления FtsZ к мембране. Другой белок, FtsA, цитоплазматический, но часто ассоциирует с мембраной. Z-кольцо может сформироваться в отсутствие одного из этих двух белков, но не может сформироваться если отсутствуют оба. По-видимому, в стабилизации Z-кольца и, возможно, в прикреплении его к мембране белки ZipA и FtsA отчасти дублируют друг друга.

Продукты нескольких других генов *fts* в установленном порядке присоединяются к Z-кольцу после включения в него FtsA. Все они — трансмембранные белки. Всю структуру Z-кольца, в ее окончательном виде, иногда называют **септальным кольцом**. Считается, что оно обладает способностью стягивать мембрану. Одним из последних в кольцо включается белок FtsW, который принадлежит к семейству SEDS. Его ген, *ftsW*, экспрессируется как единый оперон с *ftsI*, кодирующий транспептидазу; другое ее название — PBP3 (penicillin-binding protein 3). Транспептидаза FtsI, или PBP3, закреплена в мембране, а ее каталитический центр находится в периплазме (не в цитоплазме, а между двумя мембранами клеточной оболочки). Белок FtsW отвечает за включение FtsI в септальное кольцо. Образование септы можно представить в виде следующей модели: активность транспептидазы заставляет пептидогликан расти внутрь, подталкивая туда внутреннюю мембрану и увлекая за собой наружную мембрану.

Белок FtsZ является главным цитоскелетным компонентом образующейся септы. Он обычен для бактерий, а также обнаружен в хлоропластах растений. На РИС. 17.8

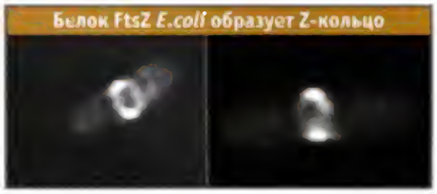


РИС. 17.7 Иммунофлюоресценция с антителами к FtsZ показывает его локализацию по средней окружности клетки. Фотография любезно предоставлена William Margolin, University of Texas Medical School at Houston



РИС. 17.8 Иммунофлюоресценция с антителами к FtsZ1 и FtsZ2 *Arabidopsis* показывает локализацию этих двух белков по средней окружности хлоропласта (фото вверху). Очертания хлоропластов четко видны в светлом поле (фото внизу). Фотография любезно предоставлена Katherine Osteryoung, Michigan State University

показана локализация белка, гомологичного FtsZ, в виде кольца по центру хлоропласта. У хлоропластов есть и другие, участвующие в делении, гены, родственные бактериальным. У бактерий и хлоропластов, как известно, общие эволюционные «корни», и их аппараты деления в целом сохранили свое сходство.

Митохондрии, которым также, вполне обоснованно, приписывают общее с бактериями происхождение, обычно не имеют белка, похожего на FtsZ. Вместо него в митохондриях используется разновидность белка динамина, который задействован в отпочковании везикул с мембран в цитоплазме эукариотической клетки. Динамин действует с наружной (цитоплазматической) стороны органелл, сжимая мембрану, чтобы на ней образовалась перетяжка.

Таким образом, общей отличительной особенностью деления у бактерий, хлоропластов и митохондрий можно назвать использование цитоскелетного белка, который образует вокруг клетки (или органеллы) кольцо, а затем или тянет, или толкает мембрану, чтобы на ней образовалась перетяжка.

другая полярная поверхность соответствует септе, возникшей в одном из предыдущих делений. Возможно, вокруг полюсов сохраняются какие-то остатки перисептальных колец, и эти остатки могут снова выступать в качестве организаторов повторного образования септы на старом месте.

Локус *minB* содержит три гена: *minC*, *minD* и *minE*. Продукты генов *minC* и *minD* работают как ингибитор деления. Точнее, MinD необходим для активации MinC; активированный MinC препятствует полимеризации FtsZ в виде Z-кольца (см. РИС. 17.9).

Экспрессия MinC/D в отсутствие MinE (или избыточная экспрессия MinC/D даже в присутствии MinE) приводит к тому, что бактерии вообще перестают делиться и вырастают в виде длинных филаментов, в которых нет септ. При экспрессии MinE, сопоставимой по уровню с экспрессией MinC/D, ингибирование деления эффективно не на всем клеточном пространстве, а только в окрестностях полюсов, так что бактерии делятся нормальным образом. Можно сказать, что белок MinE защищает центральную область клетки от ингибирования септации. Соответственно, чрезмерная экспрессия MinE индуцирует продукцию мини-клеток, поскольку

17.6 Местоположение септы регулируется генами *min*

Ключевые понятия:

- Положение, по которому образуется септа, контролируют гены *minC*, *minD* и *minE*.
- Число и расположение септ определяются соотношением концентраций белков MinE/MinC,D.
- Септа образуется там, где белку MinE удается сформировать кольцо.
- Нормальные значения концентраций MinC/D в бактериальной клетке приводят к образованию центрального кольца, но при этом подавляется образование MinE дополнительных колец у полюсов клетки.

Информацию о том, каким образом у бактерий задается местоположение септы, удалось раскрыть благодаря мутантам — продуцентам мини-клеток. Первоначально фенотип с мини-клетками был объяснен мутациями в локусе *minB*. Делеция *minB* вызывает продукцию мини-клеток, потому что в отсутствие *minB* септация становится возможной (или происходит исключительно) у полюсов клетки — в отличие от нормальной ситуации, когда септа образуется посередине клетки. Фенотип говорит о существовании у клетки подавленного стремления к образованию септ поблизости от полюсов и что роль локуса *minB* дикого типа состоит в подавлении септации у полюсов. Применив эту ситуацию к рис. 17.3, можно сказать, что у клетки, которая только что обособилась в результате деления, есть потенциал к образованию септ по трем положениям: по перисептальному кольцу и у полюсов. Один из полюсов клетки, вместе с окрестностью оболочки бактерии вокруг этого полюса, происходит от септы последнего деления;

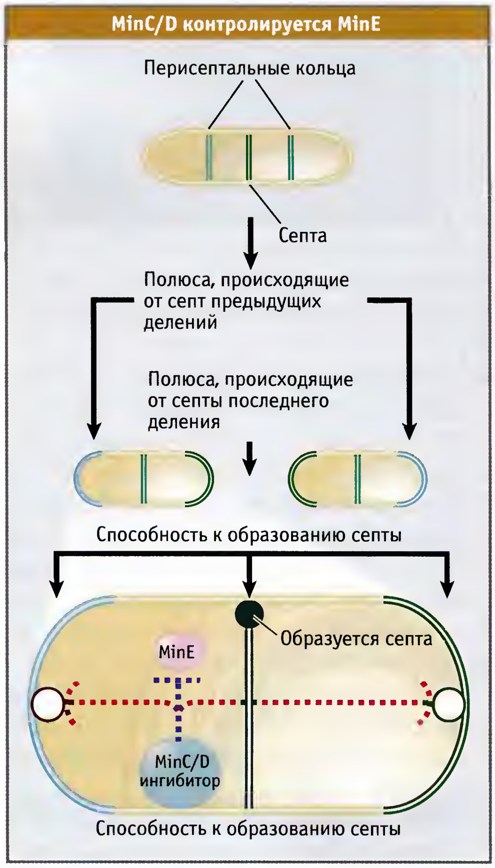


РИС. 17.9 MinC/D работает ингибитором деления; благодаря MinE он эффективен только в полярных областях бактериальной клетки

избытка MinE хватает, чтобы нейтрализовать ингибирование септации не только в центральной области клетки, но и в окрестностях полюсов, — это позволяет септам образовываться и там, и там.

Разрешающим фактором септации, таким образом, является соотношение концентраций MinC/D и MinE. При значениях этого соотношения, стандартных для бактерий дикого типа, септация у полюсов находится под запретом, а посередине клетки — разрешается. Изменения концентраций MinC/D или MinE сопровождаются взаимно-обратными эффектами. Нехватка MinC/D (или избыток MinE) приводит к «септации без разбора»; в итоге продуцируются мини-клетки. Напротив, избыток MinC/D (или нехватка MinE) ингибирует септацию повсеместно; в итоге бактерии формируют филаменты.

В том месте, где у бактерии образуется септа, белок MinE локализуется в виде кольца. Скопление MinE подавляет действие MinC/D в своей окрестности, и это работает как разрешение на приход других белков, в том числе FtsZ и ZipA, из которых формируется септальное кольцо. Интересно, что для образования кольцевого скопления MinE требуется присутствие MinD.

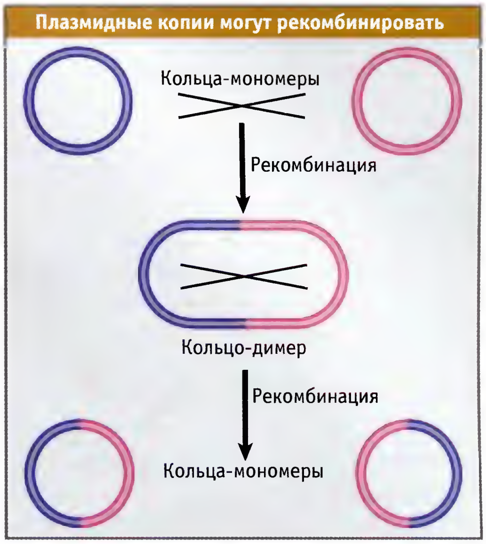


РИС. 17.10 Межмолекулярная рекомбинация сливает мономеры в димеры; внутримолекулярная рекомбинация приводит к высвобождению мономеров из состава олигомеров

17.7 Для сегрегации хромосом может быть необходима сайт-специфичная рекомбинация

Ключевые понятия:

- Система Xer осуществляет сайт-специфичную рекомбинацию последовательности-мишени рядом с участком терминации репликации хромосомы для того, чтобы обособить бактериальные хромосомы, превращенные в ходе гомологичной рекомбинации в димер.

Если в одной бактерии присутствует много копий одной плазмиды, то эти копии могут рекомбинировать между собой (поскольку последовательности их ДНК идентичны). Последствия представлены на РИС. 17.10. Одиночное событие рекомбинации между двумя кольцевыми молекулами ДНК порождает кольцо-димер; повторные события рекомбинации с участием таких димеров могут порождать кольца-мультимеры. В целом события такого рода приводят к уменьшению числа физических единиц сегрегации. В экстремальном случае с однокопийной плазмидой, которая только что реплицировалась, образование димера посредством рекомбинации означает, что теперь у клетки есть всего одна единица сегрегации (этот самый димер), так что одной из дочерних клеток плазмиды попросту не достанется. Для борьбы с такими эффектами у многих плазмид предусмотрены системы **сайт-специфичной рекомбинации**. Под действием таких систем, с участием конкретных нуклеотидных последовательностей-мишеней, димер плазмиды претерпевает внутримолекулярную рекомбинацию и превращается в два обособленных мономера.

Такие же события могут происходить и с хромосомами самих бактерий; их влияние на сегрегацию, показано на РИС. 17.11. Если рекомбинации не было, никакой проблемы не возникнет, и обособленные дочерние хромосомы беспрепятственно сегрегируют по дочерним клеткам. Если же, вслед за раундом репликации, между дочерними хромосомами произойдет гомологичная рекомбинация, то образуется димер. В этом случае дочерние хромосомы не могут обособиться; чтобы их разъединить, нужна вторая рекомбинация, аналогичная той, что происходит с плазмидными димерами.

У большинства бактерий с кольцевой хромосомой есть система сайт-специфичной рекомбинации Xer. У *E. coli* система Xer состоит из двух рекомбиназ, XerC и XerD, действующих на сайт-мишень — последовательность длиной 28 пн. Этот сайт, *dif*, расположен на хромосоме поблизости от сайта терминации репликации. Использование системы Xer клеткой соотнесено с делением этой клетки интересным образом; схема событий приведена на РИС. 17.12. Белок XerC может связаться с парой последовательностей *dif* и образовать структуру Холлидея между ними. Возможно, комплекс формируется вскоре после того, как сайт *dif* будет пройден репликативной вилкой: это объясняло бы надежность, с которой две копии сайта-мишени находят друг друга. Разбивка структуры Холлидея, приводящая к размежеванию рекомбинантов, происходит только в присутствии белка FtsK, который локализуется в септе и является необходимым для сегрегации хромосом и деления клетки. Помимо всего прочего, сайт-мишень *dif* должен

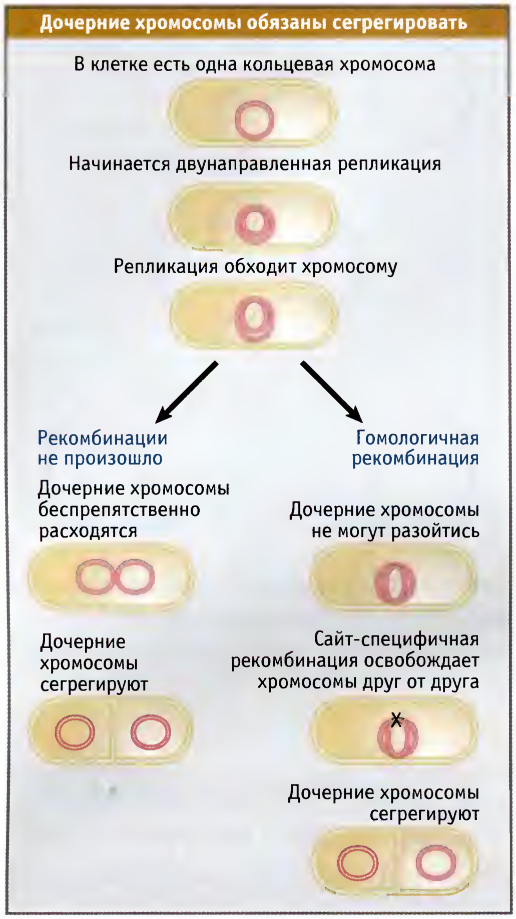


РИС. 17.11 Обычно кольцевая матричная хромосома реплицируется, образуя две дочерних хромосомы-мономера, которые сегрегируют по дочерним клеткам. Но иногда во время репликации дочерние хромосомы успевают рекомбинировать и образовать димер. Его приходится разбивать на мономеры посредством сайт-специфичной рекомбинации

располагаться на хромосоме обязательно внутри конкретной области в 30 тпн; если *dif* искусственно вынести за пределы этой области, он не сможет поддержать реакцию.

Таким образом, опция с сайт-специфичной рекомбинацией оказывается доступна, когда последовательность участка терминации репликации находится поблизости от септы. Воспользоваться этой опцией бактерия сможет, только если у нее есть димер, порожденный событием гомологичной рекомбинации. (Если такого димера нет, то сайт-специфичная рекомбинация между двумя мономерами сама породит димер!) Каким образом система Xer узнает о состоянии дочерних хромосом: существуют ли они сами по себе, или гомологичная рекомбинация объединила их в димер?

Помочь ответить на этот вопрос может то обстоятельство, что процесс сегрегации хромосом начинается вскоре после репликации. Если рекомбинации не произошло, две хромосомы отходят друг от друга. Но если все-таки образовался димер, то некоторые хромосомные последовательности, причастные к разделению таких димеров, будут ограничены в своих возможностях отдалиться друг от друга и принуждены оставаться в окрестностях септы, где на них вскоре начнет воздействовать система Xer.

У бактерий, обладающих системой Xer, всегда есть гомолог FtsK, и наоборот; таким образом, разбивка димера в около-септальном клеточном компартменте — это эволюционно сложившаяся ситуация. FtsK, крупный трансмембранный белок, привязан к септе за N-концевой домен, ассоциированный с мембраной. C-концевой домен FtsK несет две функциональных нагрузки. Одна из них — это заставлять систему Xer разделять хромосомный димер на мономеры. Кроме этого, FtsK обладает АТФазной активностью, которая может быть использована этим белком для перемещения вдоль ДНК *in vitro*. Эта вторая активность может быть направле-

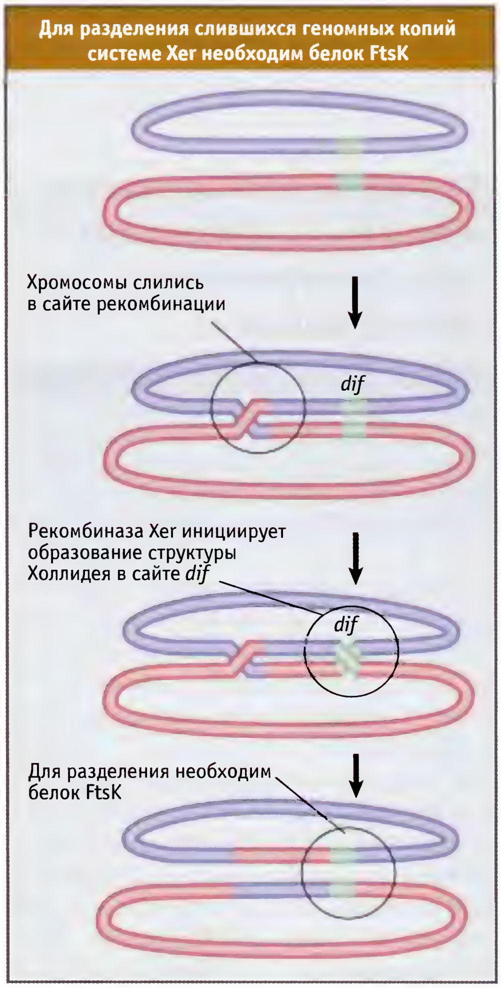


РИС. 17.12 «Несанкционированная» рекомбинация объединила геномные копии в димер. Исправляя ситуацию, система Xer образует структуру Холлидея в специальном сайте *dif*. Чтобы разбиение этой структуры дало два искомых геномных мономера, необходим белок FtsK

на на прокачку ДНК сквозь септу наподобие того, как другой белок, SpoIIIE, транспортирует ДНК в предспору при споруляции (см. разд. 17.8 «Деление вовлекает сепарацию хромосом»).

17.8 Деление вовлекает сепарацию хромосом

Ключевые понятия:

- Ориджины репликонов, возможно, прикреплены ко внутренней клеточной мембране.
- При делении хромосомы быстро сдвигаются из положения посередине клетки в положения $^1/_4$ и $^3/_4$.

Во время деления дочерние хромосомы занимают позицию по разным сторонам септы. Для правильного распределения (сегрегации) во время деления этим двум хромосомам необходимо:

- Во-первых, освободиться друг от друга настолько, чтобы расстаться сразу после терминации репликации. Для этого надо устранить защепление молекул ДНК в окрестности участка терминации. Обстоятельства таковы, что к моменту завершения репликации дочерние молекулы ДНК нуждаются в распутывании — об этом говорит тот факт, что большинство мутаций, при которых нарушена сегрегация, приходится на гены, кодирующие топоизомеразы — ферменты, умеющие пропускать цепи ДНК друг сквозь друга. У соответствующих мутантов генетический материал располагается посередине клетки одним большим комком, и образование септы приводит к обособлению безъядерной клетки от клетки, где содержатся обе дочерние хромосомы. Таким образом, чтобы успешно размножаться, бактерия обязана решать топологическую задачу распутывания реплицированных хромосом.
- Мутации, нарушающие сам процесс сегрегации, встречаются довольно редко. Априори их можно разделить на две группы: *цис*-действующие мутации в регуляторных последовательностях-мишенях и *транс*-действующие мутации в генах, чьи продукты обслуживают сегрегацию. Этими продуктами могут быть ДНК-связывающие белки или ферменты, локализованные в потенциальных местах прикрепления ДНК к оболочке бактерии. И те, и другие мутации были идентифицированы только в системах, обеспечивающих сегрегацию плазмид, хотя несколько мутаций второй группы (*транс*-действующих) было также обнаружено в бактериальных хромосомах. Сегрегацию плазмид также могут нарушать мутации в плазмидных системах сайт-специфичной рекомбинации. Эти мутации учащают потери плазмид, поскольку из-за них объектом сегрегации становится один плазмидный димер, и по фенотипическому проявлению их можно отнести к мутациям, нарушающим сегрегацию.

В своей первоначальной версии, изображенной на рис. 17.8, модель сегрегации хромосом подразумевала, что материал оболочки бактериальной клетки внедряется между сайтами прикрепления дочерних хромосом, расталкивая их в разные стороны. На самом же деле клеточная стенка и мембрана растут, если рассматривать всю поверхность клетки, непропорционально. К тому же, реплицированные хромосомы способны быстро перемещаться на новые позиции (по одной и трем четвертям клеточной длины). Если ингибировать синтез белка *до* того, как закончится репликация, хромосомы не смогут сегрегировать и сохранят свою позицию посередине клетки. Возобновление синтеза белка заставит хромосомы двинуться из середины на позиции по четвертям — хотя сама клетка при этом удлиняться не будет. Следовательно, перемещение хромосом на новые позиции, соответствующие этапу клеточного деления, является активным процессом (хотя бы потому, что этот процесс требует синтеза белка).

Нарушение сегрегации — типичный эффект мутаций в генах *muk*. В штаммах-носителях таких мутаций безъядерное потомство появляется во много раз чаще, чем обычно: вместо того, чтобы сегрегировать, дочерние хромосомы остаются по одну сторону септы. Мутации в генах *muk* нелетальны; по-видимому, генам *muk* соответствуют некие компоненты аппарата сегрегации. Так, ген *mukA* оказался не чем иным, как геном *tolC*: он кодирует известный мембранный белок, задействованный в прикреплении хромосомы к оболочке бактерии. Ген *mukB* кодирует крупный (180 кДа) глобулярный белок, по своей организации очень похожий на эукариотические белки SMC (от *англ.* structural maintenance of chromosomes — поддержание структуры хромосом). Белки SMC участвуют в конденсации и попарном скреплении дочерних хромосом (см. разд. 31.6 «Конденсация хромосомы — функция конденсинов»). Белки, похожие на SMC, были обнаружены не только у *E. coli*, но и у других бактерий.

Понимание роли белка MukB пришло с открытием, что некоторые мутации в его гене могут быть супрессированы мутациями в *topA* — гене, кодирующем топоизомеразу I. Выяснилось, что MukB образует комплекс с двумя другими белками, MukE и MukF. Комплекс MukBEF считается аналогом конденсинов эукариот. Он уплотняет (конденсирует) хромосому по механизму, основанному на сверхспирализации. Нарушение функции MukBEF приводит к неудачной сегрегации. Дефект можно компенсировать, запретив (при помощи мутаций) топоизомеразам релаксировать отрицательные сверхвитки хромосомной ДНК: в результате плотность сверхвитков увеличится, и конденсированное состояние, необходимое для сегрегации, будет восстановлено.

Таким образом, позиционирование генетического материала в бактериальной клетке может быть связа-

но с его конденсацией. Как именно оно осуществляется, пока неясно; модель, принятая на сегодняшний день, приведена на РИС. 17.13. Генетический материал исходно занимает в клетке центральное положение. Чтобы пропустить этот материал через аппарат репликации, его приходится деконденсировать. Дочерние хромосомы, возникающие в результате репликации, распутываются под действием топоизомераз, а затем, во все еще неконденсированном состоянии, поступают в распоряжение MukBEF, который занимается их конденсацией уже в положениях, которые станут центральными после того, как клетка разделится.

Предположение о физической связи бактериальной ДНК с клеточной мембраной вынашивалось годами. Но подтверждения пока довольно косвенные. К ним относится тот факт, что при фракционировании бактериальных клеток мембраносодержащие фракции оказываются обогащены генетическими маркерами окрестностей ориджина, участка терминации и компонентов репликативной вилки, а также тот факт, что белки тех же мембраносодержащих фракций оказываются инактивированными в результате разнообразных мутаций, мешающих инициации репликации.

Во время споруляции у *Bacillus subtilis* одна из дочерних хромосом должна распределиться (сегрегировать) в небольшой предспоровый компартмент (см. рис. 11.41). Этот необычный процесс, с переносом хромосомы сквозь едва образовавшуюся септу, обусловлен белком одного из генов споруляции, *spoIIIE*. Белок SpoIIIE локализован у септы и, скорее всего, работает как насос, перекачивающий ДНК в предспоровый компартмент.

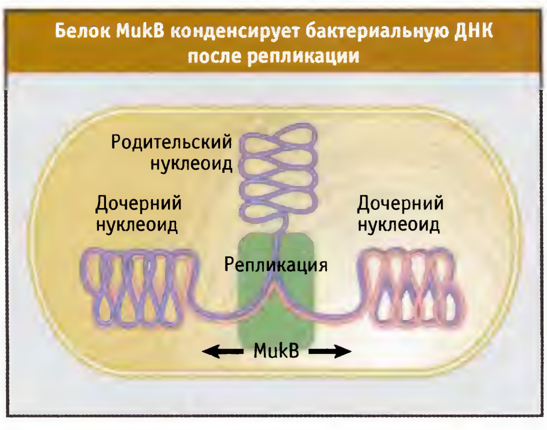


РИС. 17.13 Во время репликации ДНК бактериального нуклеоида деконденсируется. Белок MukB является важным компонентом аппарата конденсации дочерних дуплексов хромосомной ДНК бактерии и необходим для создания нуклеоидов уже на их основе

17.9 Однокопийные плазмиды обладают специальными системами сегрегации

Ключевые понятия:

- У однокопийных плазмид на каждый ориджин бактериальной хромосомы приходится одна копия плазмиды, а у мультикопийных плазмид — более, чем одна копия.
- Гомологичная рекомбинация между кольцевыми плазмидами порождает димеры и еще большие мультимеры.
- У плазмид есть системы сайт-специфичной рекомбинации, необходимые для восстановления мономеров за счет внутримолекулярной рекомбинации.
- У плазмид также есть системы сегрегации, которые гарантируют распределение дуплицированных плазмид по разным дочерним клеткам.

Каждая разновидность плазмид должна поддерживаться у бактерии-хозяина в характерном **числе копий**. Гарантированное попадание копий плазмиды в каждую из дочерних клеток обеспечивается системой, тип которой зависит от того, в каком режиме реплицируется данная конкретная плазмида:

- Системы контроля, обеспечивающие присутствие в клетке единственной копии плазмиды, похожи на аналогичную систему для бактериальной хромосомы. Они обеспечивают один раунд репликации плазмиды на каждое деление клетки. Можно сказать, что однокопийная плазмида поддерживает паритет с бактериальной хромосомой.
- Другие системы контроля допускают наличие множества событий инициации в одном клеточном цикле. В результате в каждой бактериальной клетке присутствует по нескольку копий плазмиды. Мультикопийные плазмиды характеризуются определенным числом копий, приходящихся на одну бактериальную хромосому.

Число плазмидных копий зависит в первую очередь от механизма контроля репликации. Система, ответственная за инициацию репликации, определяет, сколько плазмидных ориджинов может присутствовать в бактерии. Поскольку каждая плазмида представляет собой одиночный репликон, число ориджинов равняется числу плазмидных молекул.

У однокопийных плазмид есть система контроля репликации, по своему эффекту аналогичная системе контроля репликации бактериальной хромосомы. Единственный плазмидный ориджин может быть реплицирован только один раз, после чего дочерние ориджины сегрегируют по разным дочерним клеткам.

У многокопийных плазмид, напротив, система контроля репликации устроена так, что допускает существование целого пула ориджинов. Если число ориджинов само по себе достаточно велико (более 10 на одну бактерию), система активной сегрегации становится попросту не нужна, поскольку даже случайное распределение плазмидных копий по дочерним клеткам будет приводить к потере плазмиды с вероятностью не более 10^{-6}.

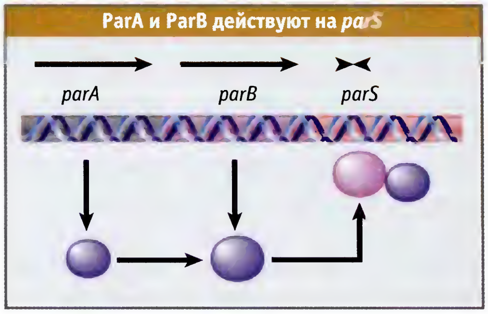


РИС. 17.14 Система сегрегации, присущая многим однокопийным плазмидам, состоит из двух генов, *parA* и *parB*, и элемента *parS*

В популяциях бактерий плазмиды всегда поддерживают себя очень аккуратно: потеря плазмиды обычно случается не чаще, чем один раз на 10^7 клеточных делений (даже для однокопийных плазмид). Системы, контролирующие сегрегацию плазмид, идентифицируются мутациями, из-за которых частота потерь плазмиды повышается, несмотря на то что репликация продолжает идти без каких-либо нарушений.

Чтобы гарантировать себе устойчивое присутствие в бактериальной популяции, плазмида может пользоваться механизмами сразу нескольких типов. Плазмидам свойственно нести в себе информацию о нескольких системах, нередко относящихся к различным типам, совместные усилия которых направлены на то, чтобы гарантировать плазмиде выживание. Некоторые из этих систем действуют окольными путями, в то время как другие системы регулируют непосредственно событие разделения. Все же те и другие системы служат одной цели: гарантировать передачу плазмиды возможно большему числу бактерий-потомков.

Однокопийные плазмиды нуждаются в системах разделения. Системы разделения следят, чтобы во время деления копии-дубликаты оказались по разные стороны септы и, таким образом, попали в разные дочерние клетки. Энзиматические активности, участвующие в разделении кольцевых молекул ДНК, впервые были описаны как раз для плазмид. Типичный пример такой системы изображен на РИС. 17.14. Система содержит два локуса, действующих в *транс*-манере, *parA* и *parB*, а также *цис*-действующий элемент, *parS*, расположенный в ДНК чуть ниже этих генов.

Белок ParA является АТФазой. Он связывается с белком ParB; тот, в свою очередь, связывает ДНК неподалеку от *parS*. Делеция любого из трех локусов *par* препятствует надлежащему разделению дубликатов плазмиды.

Системы такого типа обнаружены у плазмид F, P1, R1. Несмотря на их принципиальное сходство, никакой гомологии между соответствующими генами (и даже между *цис*-действующими элементами вроде *parS*) у них не прослеживается.

Элемент *parS* играет в плазмиде ту же роль, что и центромера хромосомы в эукариотической клетке. Белок ParB, связавший ДНК неподалеку от *parS*, — это основа структуры, распределяющей копии плазмид по разным дочерним клеткам. С самим *parS* связывается другой белок — бактериальный IHF; он становится частью той же структуры. Комплекс ParB и IHF с *parS* получил название комплекса разделения (partition complex). В последовательности *parS* (ее длина составляет 34 пн) есть сайт для связывания с белком IHF. Последовательность *parS* фланкирована последовательностями *boxA* и *boxB* — с ними связывается белок ParB.

Белок IHF, «хозяйский фактор интеграции» (integration host factor), получил свое название за ту роль, в которой он был замечен впервые. Эта роль — образование крупной структуры, обусловливающей интеграцию фага лямбда в хромосому бактерии-хозяина. Данная структура состоит из белков, вокруг которых обернута ДНК; гетеродимер IHF, как было сказано, играет ключевую роль в ее образовании. В случае с разделением плазмидных копий, роль IHF, аналогично, заключается в сгибании ДНК таким образом, что белку ParB удается связать сразу два взаимноудаленных сайта, в *boxA* и в *boxB* (показано на РИС. 17.15). Образование комплекса разделения инициируется, когда IHF связывает *parS* в присутствии димера ParB. При этом у других димеров ParB появляется возможность связать ДНК в кооперативной манере. Получившаяся структура обязана вступить в контакт с белком ParA: это взаимодействие, хоть оно и кратковременное, очень существенно для нормальной функции комплекса разделения.

Как известно, ДНК-белковый комплекс, монтируемый на IHF в процессе интеграции фага лямбда, связывает две молекулы ДНК — чтобы они смогли рекомбинировать (см. разд. 19.19 «Рекомбинация фага лямбда происходит в интасоме»). Комплекс разделения плазмидных копий, где «правит» тот же IHF, играет обратную по смыслу роль — гарантирует расставание двух молекул ДНК друг с другом; конкретного механизма обеспечения этой гарантии мы пока не знаем. Возможно, комплекс разделения физически локализует ДНК в каком-то определенном месте. Если, например, комплекс разделения прикрепляет ДНК к клеточной мембране, то задача по

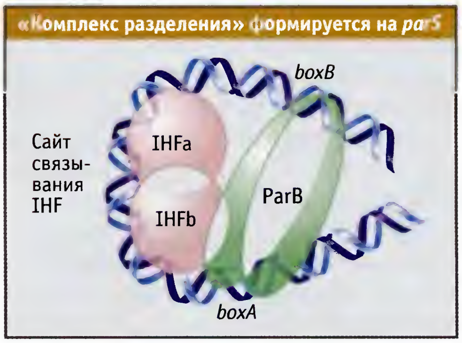


РИС. 17.15 Комплекс разделения образуется после того, как IHF свяжет *parS*. От этого связывания ДНК согнется так, что подошедший белок ParB сможет связать на ней сайты, расположенные на флангах *parS*. Первичный комплекс гетеродимера IHF и одного димера ParB дополняется последующими димерами ParB

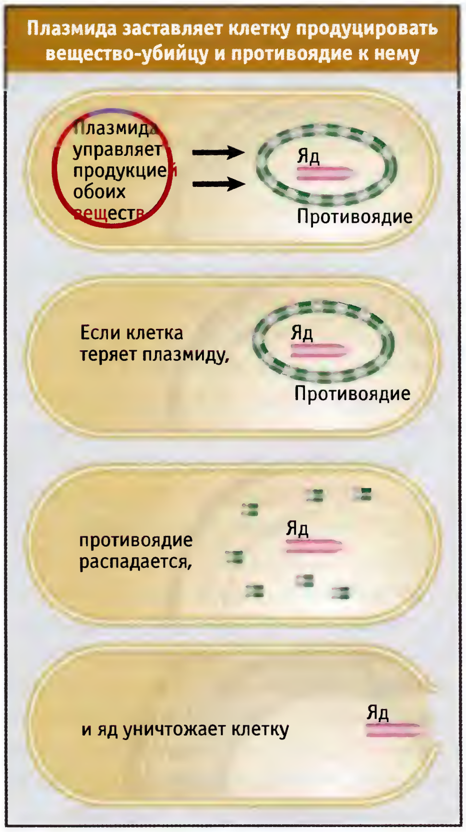


РИС. 17.16 **Некоторые плазмиды заставляют клетку-хозяина продуцировать долгоиграющий яд и короткоживущее противоядие к нему. Такая плазмида «мстит» бывшему хозяину гибелью: в результате популяция бактерий начинает состоять исключительно из клеток, в которых есть данная плазмида**

сегрегации решается образованием септы между точками прикрепления ДНК.

У нескольких бактерий были обнаружены собственные белки, родственные ParA и ParB. У *B. subtilis* эти белки получили названия Soj и SpoOJ соответственно. Мутации по их локусам препятствуют споруляции, поскольку нарушают сегрегацию одной из дочерних хромосом в предспоровый компартмент (см. рис. 11.42). В нормальных спорулирующих клетках белок SpoOJ локализован у полюса: возможно, он удерживает там ориджин. Последовательность ДНК, с которой связывается SpoOJ, во множестве копий разбросана по хромосоме *B. subtilis* — вернее, по окрестностям ориджина репликации на интервале, составляющем ~20% длины хромосомы. Скорее всего, SpoOJ связывает и старые, и новосинтезированные ориджины до тех пор, пока хромосомы не разойдутся к противоположным полюсам.

У *Caulobacter crescentus* белки ParA и ParB находятся у полюсов бактерии, где ParB связывает последовательности окрестностей ориджина, удерживая его в приполярной области цитоплазмы. Значит, у *C. crescentus* есть специальный аппарат, ответственный за локализацию ориджина у полюса клетки, и следующим этапом анализа этой ситуации станет идентификация клеточных компонентов, с которыми взаимодействует этот аппарат.

Как уже говорилось, большинство плазмид обладает несколькими независимыми системами, гарантирующими правильное разделение плазмидных копий. Этим подчеркивается важность таких систем для выживания плазмиды. Системы **жизненной зависимости** работают по принципу «живем вместе, или порознь умрем!». Такая система гарантирует, что бактерия, которая несет плазмиду, будет жить только до тех пор, пока эту плазмиду не потеряет. Существует несколько надежных способов убить клетку за то, что она избавилась от плазмиды. Эти способы объединяет общий принцип, показанный на РИС. 17.16. Его суть в том, что плазмида-резидент навязывает клетке и яд, и противоядие к нему одновременно. Яд — это вещество-убийца, которое довольно стабильно. Противоядие, блокирующее действие яда, напротив, не стабильно: оно быстро метаболизируется.

Предположим, клетка потеряла плазмиду. Противоядие, по своей природе, быстро распадается, и тогда вещество-убийца беспрепятственно приводит клетку к гибели. Таким образом, бактерии, потерявшие плазмиду, неизбежно погибают. Популяция обречена на бесконечно долгое сохранение в себе такой плазмиды.

По своей конкретной форме системы жизненной зависимости бывают разные. Например, у плазмиды F убийца и его блокатор представляют собой белки. У плазмиды R1 еще интереснее: экспрессия токсичного белка-убийцы подавляется короткой антисмысловой РНК, выступающей в качестве противоядия.

17.10 Несовместимость плазмид обусловлена сходством их репликонов

Ключевые понятия:

- **Плазмиды, принадлежащие к одной группе совместимости, имеют ориджины, регулируемые одной и той же системой контроля.**

Несовместимость плазмид — эффект работы систем регуляции числа плазмидных копий и их разделения. **Группой совместимости** называют набор плазмид, из которых любые две не могут сосуществовать в одной бактериальной клетке. Несовместимость плазмид обусловлена их сходством. Плазмиды одной группы совместимости неотличимы друг от друга — с точки зрения некоторых систем, ответственных за их поддержание. Влияние этих систем актуально на одной из двух стадий: репликации ДНК либо сегрегации генетического материала.

Применительно к несовместимости плазмид, модель негативного контроля строится в соответствии с представлением о присутствии в клетке репрессора, который выступает мерилом концентрации ориджинов. Контроль числа плазмидных копий осуществляется благода-



РИС. 17.17 Две плазмиды несовместимы (то есть принадлежат к одной группе совместимости), если на стадии инициации репликации их ориджины неотличимы друг от друга (с точки зрения клеточных систем, которых это касается). Тот же принцип применим к моделированию сегрегации хромосом

ря синтезу в клетке такого репрессора; формально, эта модель эквивалентна модели титрования, описывающей регуляцию репликации бактериальной хромосомы.

Ввод в систему нового ориджина (то есть второй плазмиды той же группы) имитирует результат репликации плазмиды-резидента: налицо присутствие двух ориджинов. Из-за этого любая дальнейшая репликация будет находиться под запретом до тех пор, пока две плазмиды не распределятся по разным клеткам. Тогда в каждой из этих дочерних клеток будет восстановлено число копий, приемлемое в качестве «числа копий до репликации» (см. РИС. 17.17).

Если две разные плазмиды не сможет различить другая система, контролирующая собственно распределение (сегрегацию, разделение) плазмид по дочерним клеткам, эффект будет тем же. Например, если две плазмиды обладают одинаковыми *цис*-действующими сайтами для разделения, конкуренция между этими сайтами будет гарантировать сегрегацию этих плазмид по разным клеткам; как следствие, они не смогут быть удержаны одной клеточной линией.

Напротив, присутствие плазмиды из одной группы совместимости не влияет на успех размножения плазмиды, принадлежащей к другой группе (во всяком случае, напрямую не влияет). На каждую группу совместимости однокопийных плазмид в одной бактериальной клетке может приходиться только один репликон, и он никак не взаимодействует с репликонами других групп совместимости.

17.11 Система совместимости ColE1 контролируется РНК-регулятором

Ключевые понятия:

- Для репликации плазмиды ColE1 необходимо, чтобы сначала через ее ориджин прошла транскрипция; продукт транскрипции расщепляется РНКазой H с образованием затравочного конца.
- Регуляторная РНК I представляет собой короткую антисмысловую РНК; она взаимодействует с транскриптом и препятствует расщеплению, порождающему затравочный конец.
- Белок Rom усиливает взаимодействие между РНК I и транскриптом.

Лучше всего изучена система регуляции числа копий (сопряженная с эффектом несовместимости) плазмиды ColE1. Это мультикопийная плазмида со стабильным уровнем присутствия ~20 копий на клетку *E. coli*. Поддержание числа плазмидных копий подчинено механизму инициации репликации в ориджине ColE1 (см. РИС. 17.18).

Репликация ColE1 начинается с транскрипции: синтез РНК стартует на 555 пн выше ориджина и прочитывает ориджин насквозь. Дальше в процесс включается фермент РНКаза H; в его названии отражена специфичность к субстрату (РНК), находящемуся в гибриде с ДНК.

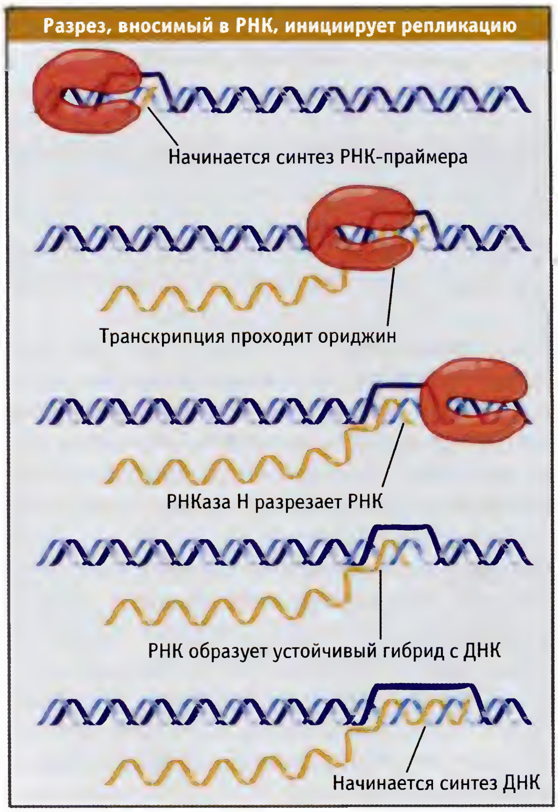


РИС. 17.18 Репликация ДНК ColE1 инициируется разрезом, который РНКаза H вносит в транскрипт-предшественник. Получается праймер со свободным 3′-OH затравочным концом, с которого начинается синтез ДНК. Праймер образует устойчивый гибрид с матричной цепью ДНК ориджина

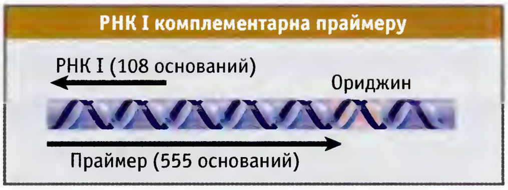


РИС. 17.19 **РНК I комплементарна 5′-концевой области праймера**

РНКаза Н разрезает транскрипт — в той части, которая комплементарна ориджину. Край этого разреза, имеющий 3′-OH конец, далее используется в качестве затравки, или праймера, для инициации синтеза ДНК. (Использование праймеров подробно обсуждается в разд. 18.8 «Чтобы начать синтез ДНК, необходима затравка».) Затравляющая РНК, или праймер, служит эффектором сразу для двух регуляторных систем. Одна из этих систем идет по пути синтеза РНК, комплементарной праймеру; другая система задействует белок — продукт близлежащего локуса. Затравляющая РНК образует с ДНК устойчивый гибрид. Участки ДНК, образующие дуплекс с РНК, расположены выше ориджина (вокруг позиций –20 и –265).

Познакомимся с еще одной молекулой: РНК I. Это особый вид регуляторной РНК, имеющий длину ~108 оснований и считываемой с другой, противолежащей цепи ДНК, чем молекула затравляющей РНК.

Взаимодействие РНК I и РНК-предшественника праймера проиллюстрировано на РИС. 17.19. Синтез молекулы РНК I инициируется внутри участка ДНК, соответствующего праймеру, и терминируется немного не доходя до места, где инициируется синтез РНК-предшественника праймера. Таким образом, РНК I комплементарна 5′-концевому участку праймера. Взаимодействие двух РНК по принципу комплементарности служит средством ограничения допуска праймера к участию в инициации раунда репликации.

Контртранскрипт — это молекула РНК, которая, как РНК I, выполняет регуляторную функцию за счет комплементарности другой РНК, считанной с той же самой области двухцепочечной ДНК. Этот тип регуляции служит еще одним примером использования антисмысловой РНК (см. разд. 13.7 «Регуляция трансляции с помощью малых молекул РНК»).

Мутации, уменьшающие или устраняющие несовместимость между плазмидами, время от времени возникают сами; их можно закрепить селекцией и картировать. У ColE1 мутации, нейтрализующие несовместимость, приходятся на участок перекрывания между РНК I и праймером. Поскольку соответствующий участок есть в обеих РНК, отношение к эффекту мутаций может иметь любая из них (или сразу обе).

Когда РНК I добавляют к системе *in vitro*, реплицирующей ДНК ColE1, она, не подавляя синтеза РНК-предшественника праймера (ни инициации транскрипции, ни элонгации), тем не менее, подавляет образование активного праймера. Следует предположить, что РНК I препятствует образованию затравочного 3′-конца, на который, как уже говорилось, воздействует РНКаза Н. Как выяснилось, присутствие РНК I мешает РНКазе Н работать. Этот ингибирующий эффект обусловлен комплементарностью последовательностей РНК I и РНК-предшественника праймера.

Обе молекулы РНК обладают одной и той же потенциальной вторичной структурой критического участка, содержащего презумптивный затравочный конец: три шпильки с петлями на концах. Мутации, смягчающие несовместимость, сосредоточены в петлях. Скорее всего, начальным этапом взаимодействия РНК I с комплементарной РНК-предшественником праймера является контакт между петлями.

Почему спаривание РНК-предшественника с РНК I мешает проведению разреза с образованием затравочного конца? Объяснение изображено на РИС. 17.20.

Если РНК I отсутствует, РНК-предшественник праймера формирует собственную вторичную структуру (со шпильками и петлями). Если РНК I присутствует, предшественник с ней взаимодействует: молекулы образуют совершенный РНК-РНК дуплекс на всем протяжении РНК I. При этом вторичная структура предшественника



РИС. 17.20 **Взаимодействие с комплементарной РНК I может менять вторичную структуру транскрипта-предшественника праймера и, таким образом, препятствовать разрезу, инициирующему синтез ДНК**

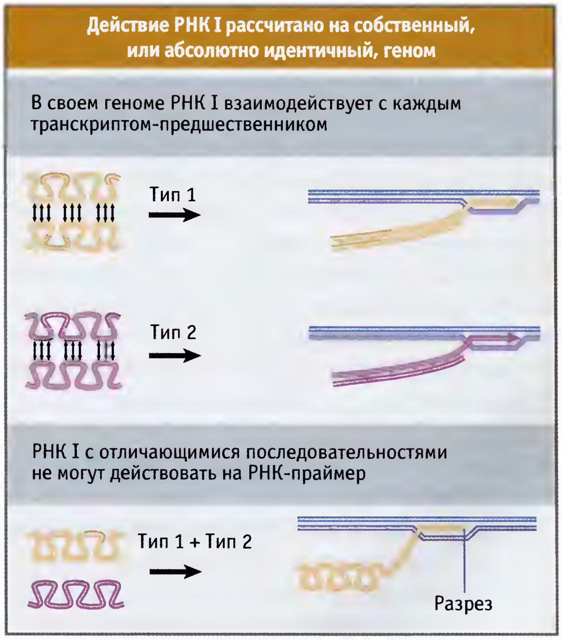


РИС. 17.21 Мутации в области, которой соответствуют РНК I и предшественник праймера, не влияют на взаимодействие этих двух РНК – продуктов одного генома. Но эти мутации могут лишить РНК I влияния на другую, прежде идентичную, плазмиду и тем самым, нарушить целостность группы совместимости

не является благоприятной для расщепления РНКазой Н: скорее всего, нарушена способность критического участка РНК к образованию устойчивого гибрида с ДНК.

Схема в целом напоминает механизм аттенуации транскрипции, где альтернативные варианты взаимодействия оснований последовательности РНК означают разрешение либо запрет на формирование вторичной структуры, необходимой для терминации транскрипции РНК-полимеразой (см. разд. 13.2 «Аттенуация контролируется с помощью альтернативных вторичных структур»). Заметим, что эффект, производимый РНК I, достигается за счет доступа к отдаленным участкам последовательности комплементарной РНК-предшественнику праймера.

Формально эту модель можно отождествить с регуляторным контуром, в котором всего два звена. Длинная РНК-предшественник праймера эквивалентна позитивному регулятору и нужна для инициации репликации. Короткая РНК I служит негативным регулятором, способным ингибировать действие позитивного регулятора.

По своей способности действовать на разные плазмиды, вводимые в клетку, РНК I выступает в роли универсального репрессора — аналогично лизогенному репрессору лямбда (см. разд. 14.9 «Репрессор и его операторы формируют регион невосприимчивости»). Только вместо белка-репрессора, который связывает новую ДНК, здесь действует РНК-репрессор, которая связывает другую РНК — новосинтезированный предшественник праймера.

На связывание РНК I с РНК-предшественником праймера может влиять белок Rom; его кодирует ген, расположенный ниже ориджина. Белок Rom усиливает связывание РНК I с транскриптами РНК-предшественника длиной >200 оснований. Это усугубляет ингибирование образования праймера.

Каким образом мутации, видоизменяющие всего-навсего вторичную структуру двух ничего не кодирующих РНК, могут исправлять несовместимость плазмид? Для ответа воспользуемся РИС. 17.21. Там показан случай, когда клетка содержит два варианта последовательности РНК I (и РНК-предшественник праймера). Если РНК I и РНК-предшественник транскрибированы с одинаковых участков, они смогут взаимодействовать. Будучи синтезированы на матрицах разных вариантов генома, они могут не узнать друг друга, и взаимодействия не получится. Конечно, такая ситуация возникнет в случае мутации не просто на участке, общем для двух РНК, но обязательно в положении, которое особенно актуально для их взаимодействия. В этом случае каждая РНК I сможет образовывать дуплекс только с предшественником, сошедшим с точно такой же плазмиды, как она сама; поэтому две плазмиды — оригинал и мутант — будут вести себя как представительницы разных групп совместимости.

17.12 Как происходит репликация и сегрегация ДНК в митохондриях эукариот?

Ключевые понятия:

- Репликация и распределение мтДНК по дочерним митохондриям носят стохастический (случайный) характер.
- Сами митохондрии распределяются по дочерним клеткам тоже произвольным образом.

За время клеточного цикла митохондрии обязаны дуплицироваться, чтобы к тому моменту, когда клетка начнет делиться, более или менее равномерно распределиться по дочерним клеткам. Механика этих процессов в целом ясна — чего нельзя сказать об их регуляции.

Размножение митохондрий в эукариотической клетке на всех своих уровнях (репликация ДНК, распределение ДНК по делящимся митохондриям, распределение митохондрий в делящейся клетке) носит свободный характер. По всей видимости, к данному случаю применимо то же, подкрепленное вероятностными выкладками, рассуждение, которое уже применялось к случаю с многокопийными плазмидами бактерий. Оно звучало так: чтобы каждой дочерней клетке практически наверняка доставалась хотя бы одна плазмидная копия, таких копий должно быть более 10 (см. разд. 17.9 «Однокопийные плазмиды обладают специальными системами сегрегации»). То же верно и для митохондрий; если мтДНК содержит аллельные вариации, случайное распределение может привести к появлению клеток, в которых есть только один из аллелей.

Репликация мтДНК, скорее всего, тоже инициируется беспорядочно, поскольку никакого явного контроля того, какие конкретно копии митохондриального генома реплицируются, нет. В каждом клеточном цикле отдельные молекулы мтДНК могут реплицироваться большее число раз, чем другие. Общее число копий митохондриального генома может контролироваться по принципу «титрования массы», сформулированному для бактерий в разд. 17.2 «Репликация взаимосвязана с клеточным циклом».

Каждая митохондрия может иметь более одной копии собственного генома. Когда митохондрия делится, она опоясывает себя кольцом, которое, сжимаясь, перетягивает органеллу надвое. Это похоже на деление бактериальной клетки.

Аппарат деления митохондрий растительной клетки похож на аппарат деления бактерий. В нем используется гомолог бактериального белка FtsZ (см. разд. 17.5 «Для образования септы необходим белок FtsZ»). У митохондрий в клетках животных аппарат деления совсем другой: в нем используется белок динамин; он также участвует в отшнуровке везикул от различных мембран внутри клетки.

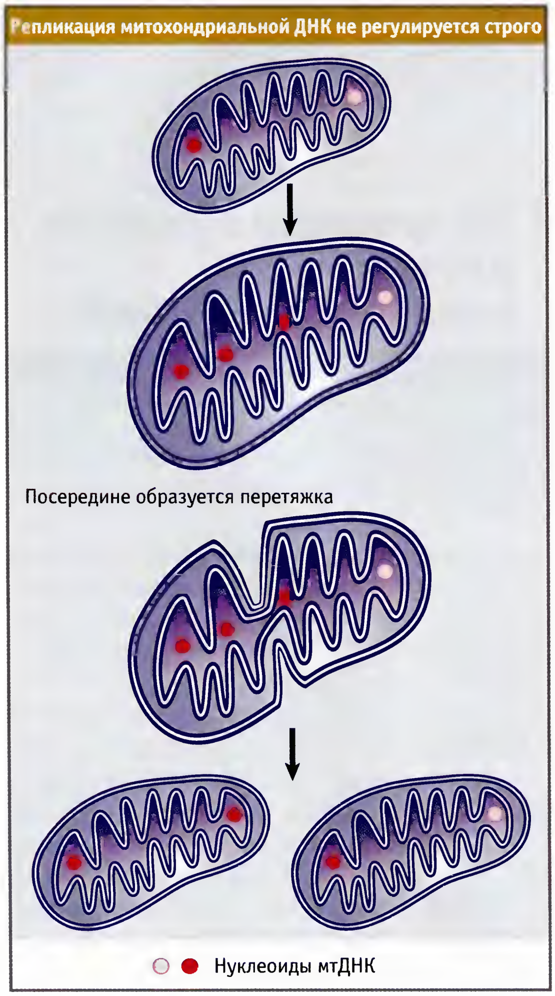


РИС. 17.22 Репликация митохондриальной ДНК регулируется: число геномных копий увеличивается пропорционально росту массы органеллы. Однако число раундов репликации, пройденных разными копиями мтДНК, никак не выровнено, и это, в совокупности с неупорядоченной сегрегацией, способствует накоплению генетических различий между митохондриями

Мы ничего не знаем о механизме разделения молекул мтДНК во внутреннем пространстве митохондрии: возможно, они достаются дочерним митохондриям только потому, что на момент деления находились в соответствующей половине органеллы. При этом возможность возникновения генетических различий между дочерними митохондриями клеткой не учитывается; между тем, как показано на РИС. 17.22, такие различия могут возникать.

Распределение митохондрий по дочерним клеткам при митозе тоже имеет случайный характер. О существовании (митохондриальных) генов, которые могут быть утрачены одной из дочерних клеток, впервые догадались, наблюдая за соматической изменчивостью у растений (см. рис. 4.15).

Размеры митохондрий даже в одной клетке могут сильно варьировать. На самом деле, до сих пор так и не выяснено: то ли каждая отдельно взятая митохондрия представляет собой дискретную органеллу, то ли является частью континуума, в котором объединены все митохондрии клетки. Континуум пребывает в непрерывном динамическом потоке слияния/обособления отдельных митохондрий.

Абсолютно точно известно, что в дрожжевых клетках митохондрии могут сливаться, поскольку спаривание двух гаплоидных штаммов дрожжей с образованием диплоидного штамма может привести к появлению рекомбинантных мтДНК. (Чтобы рекомбинировать, молекулы мтДНК должны повстречать друг друга внутри одного митохондриального компартмента, что невозможно без слияния митохондрий, унаследованных от разных штаммов.) Поиск похожих событий в клетках животных не имел явного успеха. Изучение двух клеток после слияния для обнаружения комплементации (перегруппировки) их митохондриальных аллелей не принесло однозначных экспериментальных данных.

17.13 Резюме

На репликацию хромосомы *E. coli* уходит определенное время, 40 мин; еще 20 мин должно пройти, прежде чем клетка сможет поделиться. Если клетки делятся чаще, чем раз в 60 мин, раунд репликации инициируется до окончания предыдущего клеточного цикла (клеточный цикл — это период существования клетки от деления до деления). Заблаговременная инициация раунда репликации при коротком (<60 мин) клеточном цикле приводит к появлению «хромосом с множеством вилок». Событие инициации раунда репликации регулируется путем титрования клеткой собственной массы, возможно, через накопление белка-инициатора. Скорее всего, со-

бытие инициации имеет место на клеточной мембране, поскольку в течение короткого времени после инициации ориджин ассоциирован с мембраной.

Септа, разделяющая клетку, образуется на месте перисептального кольца. Гены локуса *min* (*C*, *D* и *E*) кодируют белки-регуляторы использования перисептальных колец, основного центрального или остаточных полярных — производных от центральных колец предшествующих клеточных поколений. Если септа систематически не образуется, в колонии бактерий возникают филаменты с многочисленными ядрами; избыточное образование септ порождает безъядерные мини-клетки.

Процесс образования септы зависит от взаимодействия множества трансмембранных белков. Белок ZipA находится во внутренней мембране клеточной оболочки бактерии; он связывается с другим белком, FtsZ, который похож на тубулин и может полимеризоваться. При этом образуется волокнистая структура — Z-кольцо. С FtsZ связывается еще один белок, цитоплазматический FtsA. В дальнейшем к Z-кольцу присоединяется еще несколько продуктов *fts* (все являются трансмембранными белками). Этот процесс упорядочен, он приводит к образованию септального кольца. Последними находят свое место в септальном кольце белок семейства SEDS, называемый FtsW, и транспептидаза FtsI (или PBP3); совместно эти два белка продуцируют пептидогликаны септы. В хлоропластах и митохондриях растений используется похожий механизм деления с участием FtsZ-подобного белка; в митохондриях животных механизм деления иной: их пережимает пополам белок, похожий на динамин.

Плазмиды и бактерии обладают системами сайт-специфичной рекомбинации; эти системы решают судьбу димеров, порожденных гомологичной рекомбинацией. Разбивку таких димеров на мономеры осуществляет, например, система Xer; последовательность-мишень для этой системы расположена на участке терминации репликации хромосомы. Система Xer активна только в присутствии белка септы, FtsK; возможно, таким способом активность системы ограничивается случаями необходимости разбивки димера.

Разделение реплик плазмиды с их последующей сегрегацией по дочерним клеткам обусловлено взаимодействием белка ParB с областью *parS*, где при участии белка IHF строится специализированный комплекс разделения. Этот комплекс гарантирует сегрегацию плазмидных реплик по разным дочерним клеткам.

Механизм сегрегации может быть основан на движении ДНК: пока белок MukB конденсирует реплицированные хромосомы, их не видно. Как обособленные уплотненные массы дочерние хромосомы оформляются уже вдали друг от друга, каждая на своей половине клетки.

Плазмиды обладают множеством различных систем, обеспечивающих гарантированное разделение плазмидных копий, либо принимают в нем непосредственное участие. Число плазмидных копий характеризует присутствие конкретной плазмиды в бактериях в количестве, эквивалентном хромосомной ДНК (тогда это число равняется одной копии на единичную клетку), или в большем количестве (мультикопийно). Несовместимость двух разных плазмид в одной бактериальной клетке может быть обусловлена особенностями механизмов репликации или разделения копий. Две однокопийные плазмиды с одинаковой системой контроля репликации несовместимы друг с другом, поскольку один раунд репликации характеризует, что в геноме присутствует только одна копия плазмиды.

Литература

17.2 Репликация взаимосвязана с клеточным циклом

Обзор

Donachie, W. D. (1993). The cell cycle of *E. coli*. *Annu. Rev. Immunol.* 47, 199–230.

Исследование

Donachie, W. D. and Begg, K. J. (1970). Growth of the bacterial cell. *Nature* 227, 1220–1224.

Donachie, W. D., Begg, K. J., and Vicente, M. (1976). Cell length, cell growth and cell division. *Nature* 264, 328–333.

17.3 Септа разделяет бактериальную клетку на две дочерние, каждая из которых содержит хромосому

Обзор

de Boer, P. A. J., Cook, W. R., and Rothfield, L. I. (1990). Bacterial cell division. *Annu. Rev. Genet.* 24, 249–274.

Исследование

Spratt, B. G. (1975). Distinct penicillin binding proteins involved in the division, elongation, and shape of *E. coli* K12. *Proc. Natl. Acad. Sci. USA* 72, 2999–3003.

17.4 Мутации, нарушающие деление или сегрегацию, сказываются на форме клеток

Исследование

Adler, H. I. et al. (1967). Miniature *E. coli* cells deficient in DNA. *Proc. Natl. Acad. Sci. USA* 57, 321–326.

17.5 Для образования септы необходим белок FtsZ

Обзор

Lutkenhaus, J. and Addinall, S. G. (1997). Bacterial cell division and the Z ring. *Annu. Rev. Biochem.* 66, 93–116.

Rothfield, L., Justice, S. and Garcia-Lara, J. (1999). Bacterial cell division. *Annu. Rev. Genet.* 33, 423–438.

Исследование

Bi, E. F. and Lutkenhaus, J. (1991). FtsZ ring structure associated with division in *Escherichia coli*. *Nature* 354, 161–164.

Hale, C. A. and de Boer, P. A. (1997). Direct binding of FtsZ to ZipA, an essential component of the septal ring structure that mediates cell division in *E. coli*. *Cell* 88, 175–185.

Mercer, K. L. and Weiss, D. S. (2002). The *E. coli* cell division protein FtsW is required to recruit its cognate transpeptidase, FtsI (PBP3), to the division site. *J. Bacteriol.* 184, 904–912.

Pichoff, S. and Lutkenhaus, J. (2002). Unique and overlapping roles for ZipA and FtsA in septal ring assembly in *Escherichia coli*. *EMBO J.* 21, 685–693.
Romberg, L. and Levin, P. A. (2003). Assembly dynamics of the bacterial cell division protein FTSZ: poised at the edge of stability. *Annu. Rev. Microbiol.* 57, 125–154.
Stricker, J., Maddox, P., Salmon, E. D., and Erickson, H. P. (2002). Rapid assembly dynamics of the *Escherichia coli* FtsZ-ring demonstrated by fluorescence recovery after photobleaching. *Proc. Natl. Acad. Sci. USA* 99, 3171–3175.

17.6 Местоположение септы регулируется генами *min*

Исследование

de Boer, P. A. J. et al. (1989). A division inhibitor and a topological specificity factor coded for by the minicell locus determine proper placement of the division septum in *E. coli*. *Cell* 56, 641–649.
Hu, Z., Mukherjee, A., Pichoff, S., and Lutkenhaus, J. (1999). The MinC component of the division site selection system in *E. coli* interacts with FtsZ to prevent polymerization. *Proc. Natl. Acad. Sci. USA* 96, 14819–14824.
Pichoff, S. and Lutkenhaus, J. (2001). *Escherichia coli* division inhibitor MinCD blocks septation by preventing Z-ring formation. *J. Bacteriol.* 183, 6630–6635.
Raskin, D. M. and de Boer, P. A. J. (1997). The MinE ring: an FtsZ-independent cell structure requires for selection of the correct division site in *E. coli*. *Cell* 91, 685–694.

17.7 Для сегрегации хромосом может быть необходима сайт-специфичная рекомбинация

Исследование

Aussel, L., Barre, F. X., Aroyo, M., Stasiak, A., Stasiak, A. Z., and Sherratt, D. (2002). FtsK is a DNA motor protein that activates chromosome dimer resolution by switching the catalytic state of the XerC and XerD recombinases. *Cell* 108, 195–205.
Barre, F. X., Aroyo, M., Colloms, S. D., Helfrich, A., Cornet, F., and Sherratt, D. J. (2000). FtsK functions in the processing of a Holliday junction intermediate during bacterial chromosome segregation. *Genes Dev.* 14, 2976–2988.
Blakely, G., May, G., McCulloch, R., Arciszewska, L. K., Burke, M., Lovett, S. T., and Sherratt, D. J. (1993). Two related recombinases are required for site-specific recombination at dif and cer in *E. coli* K12. *Cell* 75, 351–361.

17.8 Деление вовлекает сепарацию хромосом

Обзор

Draper, G. C. and Gober, J. W. (2002). Bacterial chromosome segregation. *Annu. Rev. Microbiol.* 56, 567–597.
Errington, J., Bath, J., and Wu, L. J. (2001). DNA transport in bacteria. *Nat. Rev. Mol. Cell Biol.* 2, 538–545.
Gordon, G. S. and Wright, A. (2000). DNA segregation in bacteria. *Annu. Rev. Microbiol.* 54, 681–708.
Hiraga, S. (1992). Chromosome and plasmid partition in *E. coli*. *Annu. Rev. Biochem.* 61, 283–306.
Wake, R. G. and Errington, J. (1995). Chromosome partitioning in bacteria. *Annu. Rev. Genet.* 29, 41–67.

Исследование

Case, R. B., Chang, Y. P., Smith, S. B., Gore, J., Cozzarelli, N. R., and Bustamante, C. (2004). The bacterial condensin MukBEF compacts DNA into a repetitive, stable structure. *Science* 305, 222–227.
Jacob, F., Ryter, A., and Cuzin, F. (1966). On the association between DNA and the membrane in bacteria. *Proc. Roy. Soc. Lond. B Biol. Sci.* 164, 267–348.
Sawitzke, J. A. and Austin, S. (2000). Suppression of chromosome segregation defects of *E. coli* muk mutants by mutations in topoisomerase I. *Proc. Natl. Acad. Sci. USA* 97, 1671–1676.
Wu, L. J. and Errington, J. (1997). Septal localization of the SpoIIIE chromosome partitioning protein in *B. subtilis*. *EMBO J.* 16, 2161–2169.

17.9 Однокопийные плазмиды обладают специальными системами сегрегации

Обзор

Draper, G. C. and Gober, J. W. (2002). Bacterial chromosome segregation. *Annu. Rev. Microbiol.* 56, 567–597.
Gordon, G. S. and Wright, A. (2000). DNA segregation in bacteria. *Annu. Rev. Microbiol.* 54, 681–708.
Hiraga, S. (1992). Chromosome and plasmid partition in *E. coli*. *Annu. Rev. Biochem.* 61, 283–306.

Исследование

Davis, M. A. and Austin, S. J. (1988). Recognition of the P1 plasmid centromere analog involves binding of the ParB protein and is modified by a specific host factor. *EMBO J.* 7, 1881–1888.
Funnell, B. E. (1988). Participation of *E. coli* integration host factor in the P1 plasmid partition system. *Proc. Natl. Acad. Sci. USA* 85, 6657–6661.
Mohl, D. A. and Gober, J. W. (1997). Cell cycle-dependent polar localization of chromosome partitioning proteins of *C. crescentus*. *Cell* 88, 675–684.
Surtees, J. A. and Funnell, B. E. (2001). The DNA binding domains of P1 ParB and the architecture of the P1 plasmid partition complex. *J. Biol. Chem.* 276, 12385–12394.

17.10 Несовместимость плазмид обусловлена сходством их репликонов

Обзор

Nordstrom, K. and Austin, S. J. (1989). Mechanisms that contribute to the stable segregation of plasmids. *Annu. Rev. Genet.* 23, 37–69.
Scott, J. R. (1984). Regulation of plasmid replication. *Microbiol. Rev.* 48, 1–23.

17.11 Система совместимости ColE1 контролируется РНК-регулятором

Исследование

Masukata, H. and Tomizawa, J. (1990). A mechanism of formation of a persistent hybrid between elongating РНК and template DNA. *Cell* 62, 331–338.
Tomizawa, J.-I. and Itoh, T. (1981). Plasmid ColE1 incompatibility determined by interaction of РНК with primer transcript. *Proc. Natl. Acad. Sci. USA* 78, 6096–6100.

17.12 Как происходит репликация и сегрегация ДНК в митохондриях эукариот?

Обзор

Birky, C. W. (2001). The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. *Annu. Rev. Genet.* 35, 125–148.

Репликация ДНК

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

18.1 Введение

18.2 ДНК-полимеразы — ферменты, создающие ДНК

- ДНК синтезируется как при полуконсервативной репликации, так и при репарации.
- И бактерии, и эукариотические клетки имеют по нескольку разных ДНК-полимераз.
- У бактерий одна из ДНК полимераз обеспечивает полуконсервативную репликацию, остальные участвуют в репарации.
- Ядро, митохондрии и хлоропласты эукариот содержат по одной ДНК-полимеразе, осуществляющей репликацию, прочие полимеразы помогают основному ферменту в ходе репликации, а также вовлечены в процессы репарации.

18.3 ДНК-полимеразы обладают различными нуклеазными активностями

- ДНК-полимераза I обладает уникальной 5′–3′ экзонуклеазной активностью, которая в сочетании с синтезом ДНК может быть использована для ник-трансляции.

18.4 ДНК-полимеразы контролируют точность репликации

- ДНК-полимеразы часто обладают 3′–5′ экзонуклеазной активностью, которую используют для удаления неверно подставленных оснований.
- Такая корректирующая активность повышает точность репликации примерно в 100 раз.

18.5 ДНК-полимеразы имеют общую структуру

- Многие ДНК-полимеразы имеют большую выемку, образованную тремя доменами, напоминающими руку.
- ДНК лежит поперек «ладони» в бороздке, образованной «большим пальцем» и остальными «пальцами».

18.6 Синтез ДНК является полунепрерывным

- Осуществляющая репликацию ДНК-репликаза движется непрерывно, синтезируя ведущую цепь, однако синтез отстающей цепи идет посредством создания коротких фрагментов, которые затем соединяются.

18.7 Наработка одноцепочечной ДНК для размножения фага в рамках модельной системы φX

- Для репликации необходима хеликаза — чтобы разделить цепи ДНК за счет энергии гидролиза АТФ.
- Белок SSB требуется для поддержания разделенных цепей ДНК в одноцепочечном состоянии.
- Комбинация хеликазы, SSB и белка A разделяет φX174-дуплекс на одноцепочечное кольцо и одноцепочечный линейный фрагмент.

18.8 Чтобы начать синтез ДНК, необходима затравка

- Всем ДНК-полимеразам требуется 3′-OH затравочный конец.
- Затравочный конец может быть обеспечен РНК-затравкой (праймером), одноцепочечным разрывом в ДНК или затравляющим белком.
- Для репликации ДНК специальная РНК-полимераза (праймаза) синтезирует цепь РНК, обеспечивающую затравочный конец.
- У *E.coli* начало синтеза ДНК может происходить двумя способами: в бактериальном ориджине (oriC) и в ориджине φX174.
- Затравление репликации на двухцепочечной ДНК всегда требует наличия репликазы, SSB и праймазы.
- DnaB — хеликаза, разворачивающая ДНК при репликации *E.coli*.

18.9 Холофермент ДНК-полимеразы состоит из трех субкомплексов

- Репликаза *E.coli* ДНК-полимераза III — это 900 кДа комплекс с димерной структурой.
- Каждая мономерная субъединица обладает каталитическим центром, димеризационной субъединицей и компонентом процессивности.
- Специальный белок загружает субъединицы процессивности на ДНК, и они зажимают ее в кольцо.
- На каждую цепь матрицы приходится по одному каталитическому центру.

18.10 Зажим контролирует ассоциацию минимального фермента с ДНК

- Полимераза (минимальный фермент) ведущей цепи отличается высокой процессивностью, поскольку зажим удерживает ее на ДНК.
- Зажим, ассоциированный с полимеразой отстающей цепи, диссоциирует в конце каждого фрагмента Оказаки и заново собирается для следующего фрагмента.
- Хеликаза DnaB отвечает за взаимодействие с праймазой DnaG для инициации синтеза фрагмента Оказаки.

18.11 Координирование синтеза ведущей и отстающей цепей

- Для синтеза ведущей и отстающей цепей требуются разные комплексы ферментов.

- У *E.coli* оба этих комплекса содержат одну и ту же каталитическую субъединицу (DnaE).
- У других организмов для каждой из цепей могут требоваться различные каталитические субъединицы.

18.1 Введение

Репликация двухцепочечной ДНК — это комплексное мероприятие, задействующее целый конгломерат ферментативных активностей. Различные активности вовлечены в стадии инициации, элонгации и терминации.

- Инициация включает в себя опознание ориджина репликации комплексом белков. До начала синтеза ДНК матричные цепи должны быть разделены и (временно) стабилизированы в одноцепочечном состоянии. После этой стадии в репликативной вилке может быть инициирован синтез дочерних цепей.
- Элонгация находится под контролем другого белкового комплекса (**реплисомы**). Заметим, что реплисома существует как комплекс только будучи ассоциирована с ДНК конкретной структуры, присущей репликативной вилке и не существует как независимая единица (в отличие, например, от рибосомы). По мере продвижения реплисомы вдоль ДНК матричные цепи раскручиваются, а дочерние — синтезируются.
- В конце репликона необходимо провести реакцию соединения и/или терминации. Вслед за терминацией удвоенные хромосомы должны быть отделены друг от друга, что требует манипулирования структурами ДНК высшего порядка.

Неспособность реплицировать ДНК приводит к нежизнеспособности клеток в культуре. Репликативные мутации, с которыми возможно работать, должны быть условно-летальными. В пермиссивных условиях (т. е. при обычной температуре инкубации) такие мутанты способны осуществлять репликацию, а при более высокой температуре в 42 °C — нет. Всесторонний анализ серий таких чувствительных к температуре мутантов *E. coli* позволил выявить ряд локусов, названных генами dna. По тому, как такие **dna-мутанты** ведут себя при повышении температуры до непермиссивной, можно отследить две стадии репликации, к которым они и относятся:

- Большинство принадлежит к классу **быстрого реагирования** (quick-stop mutants) — они прекращают репликацию сразу при повышении температуры. Они дефектны по компонентам аппарата репликации, как правило, ферментам, необходимым для элонгации (реже — по аппарату поддержки синтеза предшественников).
- Меньшая часть принадлежит к классу **медленного реагирования** (slow-stop mutants) — они завершают текущий раунд репликации, но не могут начать следующий. Их дефекты проявляются в ходе инициации в ориджине следующего раунда репликации.

Важную роль в идентификации компонентов репликативного аппарата сыграла принципиальная схема экспериментов под названием **комплементация *in vitro***. Система для репликации *in vitro* приготавливается на основе *dna*-мутанта и находится в условиях, при которых

продукт мутировавшего гена не активен. Экстракты из клеток дикого типа тестируют на предмет восстановления этой активности. Белок, кодируемый локусом *dna*, может быть идентифицирован и очищен как активный компонент «восстанавливающего» экстракта.

Каждый компонент реапликативного аппарата бактерий в настоящее время применяется и как биохимически чистый продукт для исследований *in vitro*, и как набор мутантов по соответствующему гену для исследований *in vivo*. Репликативные системы эукариот также подвергаются высокой степени очистки (и обычно их компоненты аналогичны бактериальным белкам), однако не для всех из них идентифицирован каждый отдельный компонент.

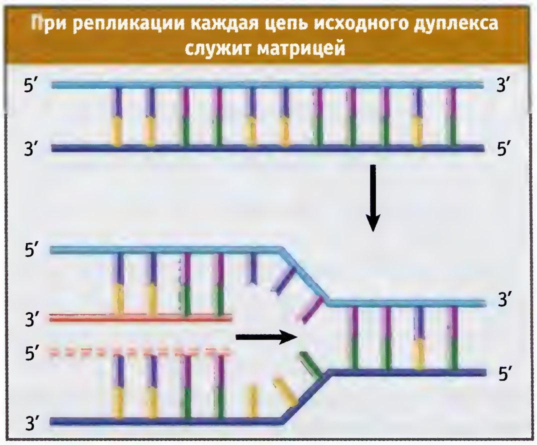


РИС. 18.1 Полуконсервативная репликация ДНК заключается в синтезе двух новых цепей на матрицах старых

18.2 ДНК-полимеразы — ферменты, создающие ДНК

Ключевые понятия:

- ДНК синтезируется как при полуконсервативной репликации, так и при репарации.
- И бактерии, и эукариотические клетки имеют по нескольку разных ДНК-полимераз.
- У бактерий одна из ДНК-полимераз обеспечивает полуконсервативную репликацию, остальные участвуют в репарации.
- Ядро, митохондрии и хлоропласты эукариот содержат по одной ДНК-полимеразе, осуществляющей репликацию, прочие полимеразы помогают основному ферменту в ходе репликации, а также вовлечены в процессы репарации.



РИС. 18.2 Репарация ДНК заключается в замене короткого участка цепи, где имело место повреждение

Существует два основных типа синтеза ДНК.

На РИС. 18.1 показан результат полуконсервативной **репликации**. Цепи дуплекса разделены, и каждая служит матрицей для синтеза новой цепи. Матричный дуплекс замещается двумя дочерними дуплексами, каждый из которых состоит из одной исходной и одной новосинтезированной цепей.

На РИС 18.2 показаны последствия **репарации**. Одна цепь ДНК была повреждена. Этот участок вырезается и замещается новосинтезированным материалом. Фермент, который может синтезировать новую цепь ДНК на ДНК-матрице, называется ДНК-зависимой **ДНК-полимеразой**. И про-, и эукариотические клетки обладают несколькими ДНК-полимеразами. Только некоторые из этих ферментов в действительности осуществляют репликацию; их иногда называют **ДНК-репликазами**. Прочие ДНК-полимеразы участвуют в репарации и/или репликации как вспомогательные.

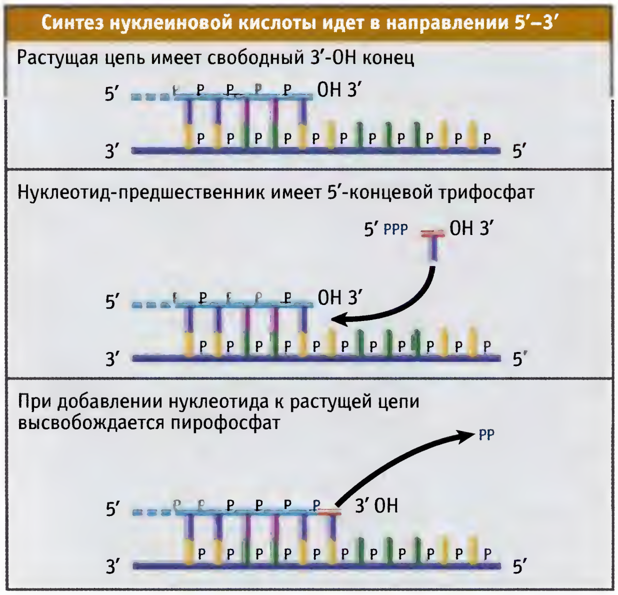


РИС. 18.3 ДНК синтезируется за счет добавления нуклеотидов к 3′-ОН-концу растущей цепи; поэтому новая цепь растет в направлении 5′-3′. Единичный предшественник представляет собой нуклеозидтрифосфат. При вступлении в синтез предшественник теряет две концевые фосфатные группы

Все про- и эукариотические ДНК-полимеразы работают по одному общему фундаментальному принципу. Они удлиняют цепь ДНК, добавляя к 3′-концу по одному нуклеотиду, выбор которого диктуется требованием комплементарности матричной цепи ДНК. Это схематически изображено на РИС. 18.3.

Некоторые ДНК-полимеразы функционируют как независимые ферменты, некоторые (а именно — репликазы)

У *E. coli* есть пять ДНК-полимераз

Фермент	Ген	Функция
I	*polA*	главный фермент репарации
II	*polB*	повторный старт репликации
III	*polC*	репликаза
IV	*dinB*	транслесионная репликация
V	*umuD'$_2$C*	транслесионная репликация

РИС. 18.4 Только одна из ДНК-полимераз *E. coli* является компонентом репликазы. Остальные ДНК-полимеразы участвуют в репарации поврежденной ДНК, перезапуске остановившихся репликативных вилок, а также в транслесионной репликации (то есть способствуют проходу репликативной вилки через поврежденный участок ДНК)

встроены в большие белковые комплексы. Собственно синтез ДНК — лишь одна из нескольких функций репликазного комплекса, который обычно обладает многими другими активностями, касающимися раскручивания ДНК, инициации синтеза новой цепи и т.д.

РИС. 18.4 дает обобщенные сведения о ДНК-полимеразах *E.coli*. ДНК-полимераза III (мультисубъединичный белок, репликаза) отвечает за синтез цепей ДНК *de novo*. ДНК-полимераза I (кодируемая геном *polA*) участвует в репарации поврежденной ДНК, а также играет вспомогательную роль в полуконсервативной репликации. ДНК-полимераза II необходима для перезапуска репликативной вилки, которая была заблокирована вследствие повреждения ДНК. ДНК полимеразы IV и V позволяют провести репликацию в обход определенных типов повреждений.

В экстрактах *E.coli*, изучаемых на предмет способности к синтезу ДНК, активность ДНК-полимеразы I доминирует, перекрывая остальные. Ее активность так велика, что детектировать работу ферментов, на самом деле ответственных за репликацию, невозможно! Поэтому системы слежения за репликацией *in vitro* готовят на основе мутантов *polA*.

Некоторые фаги кодируют собственные ДНК-полимеразы. К ним относятся T4, T5, T7 и SPO1. Полимеразы этих фагов обладают 5'–3' синтетической и 3'–5' экзонуклеазной корректирующей активностями (см. разд. 18.3 «ДНК-полимеразы обладают различными нуклеазными активностями»). Во всех случаях мутация в гене, кодирующем одиночный полипептид фаговой ДНК-полимеразы, препятствует размножению фага. Полипептид фаговой ДНК-полимеразы ассоциирует с другими белками (фага или собственно бактерии-хозяина), образуя полноценный фермент.

Идентифицировано несколько классов эукариотических ДНК-полимераз: δ и ε необходимы для репликации в ядре, α занимается праймингом (инициацией) репликации. Прочие ДНК-полимеразы участвуют в репарации ДНК (β и ε) или в репликации ДНК митохондрий (γ).

18.3 ДНК-полимеразы обладают различными нуклеазными активностями

Ключевые понятия:

- ДНК-полимераза I обладает уникальной 5'-3' экзонуклеазной активностью, которая в сочетании с синтезом ДНК может быть использована для ник-трансляции.

Вместе со способностью синтезировать ДНК репликазы часто обладают нуклеазной активностью. 3'–5' экзонуклеазная активность, как правило, используется для удаления неправильно добавленного нуклеотида, что обеспечивает систематическую коррекцию ошибок синтеза. Так репликаза корректирует сама себя (см. разд. 18.4 «ДНК-полимеразы контролируют точность репликации»).

Исторически ДНК-полимераза I *E.coli* была первым охарактеризованным ДНК-синтезирующим ферментом. Она представляет собой одиночный полипептид массой 103 кДа, который может быть расщеплен на две части посредством протеолитической обработки. Две трети белка составил С-концевой фрагмент, обладающий полимеразной активностью, оставшуюся треть — N-конец с экзонуклеазной активностью.

С-концевой фрагмент (68 кДа) получил название фрагмента Кленова. Его широко используют в лабораторной практике для синтеза ДНК *in vitro*. Он обладает полимеразной и 3'–5' экзонуклеазной активностями. Соответствующие им каталитические центры в белке разделяет расстояние около 30 Å (добавление и удаление нуклеотида разнесены в пространстве).

Меньший, N-концевой фрагмент (35 кДа) обладает 5'–3' экзонуклеазной активностью — отрезает небольшие фрагменты цепи ДНК (до 10 нуклеотидов) в один

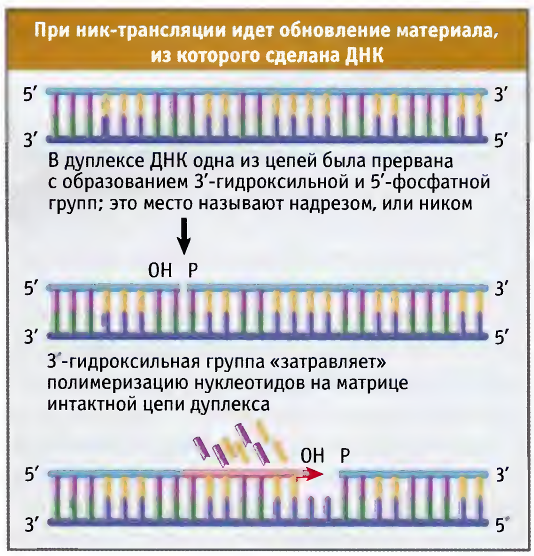


РИС. 18.5 На одном из флангов надреза идет замена материала одной из цепей. Происходит ник-трансляция: надрез движется вдоль ДНК

прием. Эта активность (в интактном полипептиде) скоординирована с синтетической и корректирующей активностями, сообщая ДНК-полимеразе I уникальную способность запускать репликацию *in vitro* с одноцепочечного разрыва — ника (nick) — в ДНК. (Ни одна другая ДНК-полимераза не способна это делать.) Начиная с точки разрыва фосфодиэфирной связи, фермент удлиняет 3′-OH конец. По мере синтеза новый сегмент растущей цепи замещает деградирующий гомологичный участок.

Этот процесс — **ник-трансляция** — изображен на РИС. 18.5. Вытесняемая цепь ДНК подвергается деградации в результате 5′–3′ экзонуклеазной активности фермента. Свойства ДНК при этом не меняются, за исключением того, что одна из цепей оказывается замещенной новосинтезированным материалом, а ник — сдвинутым в другое положение. Ник-трансляция нашла огромное практическое применение в качестве основного способа введения радиоактивно меченных нуклеотидов в ДНК *in vitro*.

Сочетание 5′–3′ синтетической и 3′–5′ экзонуклеазной активностей, вероятно, используется *in vivo* главным образом для заполнения коротких одноцепочечных участков в дуплексе ДНК. Такие «бреши» возникают в ходе репликации, а также после удаления из ДНК поврежденных оснований.

18.4 ДНК-полимеразы контролируют точность репликации

Ключевые понятия:

- ДНК-полимеразы часто обладают 3′–5′ экзонуклеазной активностью, которую используют для удаления неверно подставленных оснований.
- Такая корректирующая активность повышает точность репликации примерно в 100 раз.

Точность репликации — проблема того же рода, что и аккуратность трансляции (которую мы уже рассматривали). Точность репликации основана на специфичности спаривания азотистых оснований. Сами по себе взаимодействия, обусловливающие спаривание, с частотой 10^{-3} подвергаются ошибкам. Действительная частота ошибок репликации у бактерий составляет всего 10^{-8}–10^{-10}. Это соответствует одной ошибке на 1000 раундов репликации одного генома, или 10^{-6} ошибки в каждой новой копии гена.

Ошибки, допускаемые ДНК-полимеразой в ходе репликации, делят на два класса:

- Сдвиги рамки считывания имеют место в результате пропуска нуклеотида или вставки лишнего. Точность репликации в части защищенности от сдвигов рамки зависит от **процессивности** фермента — тенденции сохранять непрерывное сродство с субстратом (матрицей ДНК), вместо того чтобы диссоциировать или ассоциировать. Это особенно важно при репликации гомополимерных отрезков ДНК (например, длинной последовательности $dT_n{:}dA_n$), где проскальзывание репликации может изменить длину гомополимерного участка. В общем случае повышенная процессивность снижает вероятность таких ошибок. У мультимерных ДНК-полимераз высокая процессивность обычно обеспечивается отдельной субъединицей, которая не нужна для каталитической активности *per se*.
- Замены случаются, когда встроенным оказывается неправильный (неверно спаренный) нуклеотид. Частота таких ошибок определяется эффективностью **самокоррекции** (proofreading) репликазы: фермент проверяет новообразованную пару оснований и удаляет только что добавленный нуклеотид, если он спарен с матрицей неверно.

Все бактериальные ДНК-полимеразы обладают 3′–5′ экзонуклеазной активностью (в направлении, обратном ходу синтеза). В начале шага элонгации нуклеотид-предшественник занимает позицию на конце растущей цепи. Образуется связь. Затем фермент продвигается вперед еще на одну пару оснований, после чего готов принять новый нуклеотид-предшественник, и рабочий цикл повторяется. Если произошла ошибка, фермент не продвигается, а удаляет последний добавленный нуклеотид (РИС. 18.6).

Разные ДНК-полимеразы по-разному устанавливают баланс между синтетической и корректирующей активностями. В некоторых случаях эти активности «размещены» в одной белковой субъединице, в других —

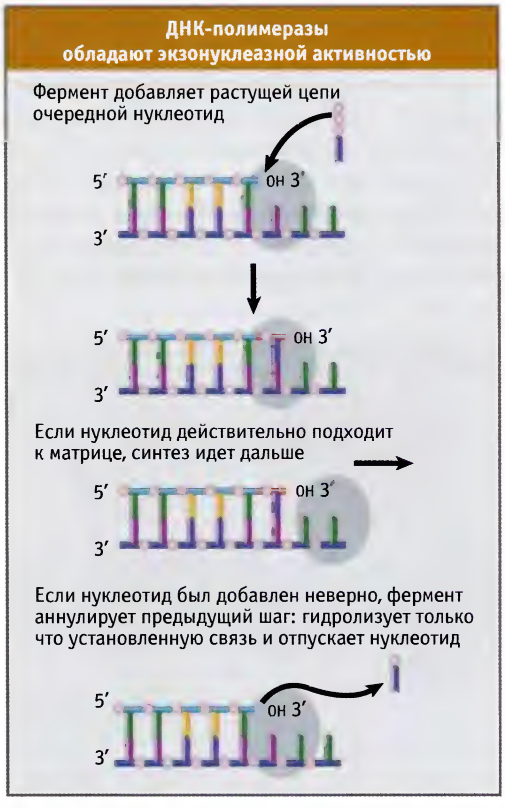


РИС. 18.6 На каждом шаге синтеза бактериальная ДНК-полимераза устраивает дополнительную проверку: насколько хорошо конец растущей цепи соответствует матрице. Если полимераза находит ошибку, она тут же удаляет только что добавленный нуклеотид

принадлежат разным. Каждая ДНК-полимераза характеризуется определенной частотой ошибок, понижаемой ее же корректирующей активностью. В типичном случае коррекция снижает частоту ошибок репликации с 10^{-5} до 10^{-7} на пару оснований. Пострепликативные системы распознавания и коррекции ошибок доводят этот показатель до $<10^{-9}$ (см. разд. 20.7 «Контроль направления репарации несовершенных пар оснований»).

Репликазная активность ДНК-полимеразы III была открыта благодаря летальной муации по локусу *dnaE*, кодирующему 130 кДа α-субъединицу с ДНК-синтезирующей активностью. 3′–5′ экзонуклеазная (корректирующая) активность обнаружена у другой субъединицы, ε, кодируемой локусом *dnaQ*. Ведущая роль субъединицы ε в контроле точности репликации *in vivo* демонстрируется фенотипом мутантов по *dnaQ*: мутации в их ДНК возникают в 10^3 раз чаще.

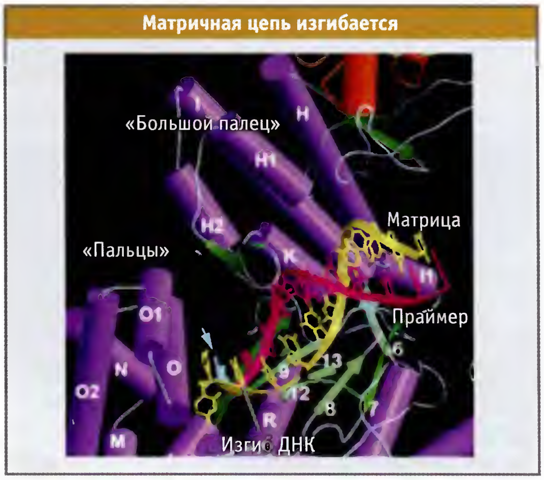


РИС. 18.8 Структурный анализ ДНК-полимеразы фага T7 показал, что матричная цепь ДНК содержит крутой изгиб, направленный навстречу входящему нуклеотиду. Фотография любезно предоставлена Charles Richardson and Thomas Ellenberger, Washington University School of Medicine

18.5 ДНК-полимеразы имеют общую структуру

Ключевые понятия:

- Многие ДНК-полимеразы имеют большую выемку, образованную тремя доменами, напоминающими руку.
- ДНК лежит поперек «ладони» в бороздке, образованной «большим пальцем» и остальными «пальцами».

На РИС. 18.7 показаны общие структурные черты всех ДНК-полимераз. В структуре фермента вычленяется несколько независимых доменов, которые в совокупности напоминают человеческую правую руку. ДНК связывается в большой выемке, образованной тремя доменами. Основу каталитического активного центра образуют консервативные аминокислотные мотивы в составе домена «ладонь».



РИС. 18.7 У ДНК-полимеразы есть «ладонь» с активным центром, «большой палец», который связывает ДНК и обеспечивает процессивность, а также экзонуклеазный домен (со своим собственным активным центром) и N-концевой домен

«Пальцы» правильным образом располагают матрицу в активном центре. «Большой палец» связывает ДНК на выходе из фермента и обусловливает высокую процессивность. В активном центре самые важные консервативные области всех трех доменов сближены и образуют непрерывную поверхность. Экзонуклеазная активность находится в независимом домене с собственным каталитическим центром. N-концевой домен внедрен в экзонуклеазный. По сходству белковой последовательности ДНК-полимеразы подразделяют на пять семейств; «ладонь» довольно консервативна, в то время как и «большой палец», и остальные «пальцы» представляют собой аналогичные структуры на основе разных аминокислотных последовательностей.

В ходе каталитической реакции в активном центре ДНК-полимеразы нуклеотидтрифосфат спаривается с цепью ДНК. Она лежит поперек «ладони» в бороздке, образованной «большим пальцем» и «пальцами». На РИС. 18.8 показана кристаллическая структура T7 ДНК-полимеразы в комплексе с ДНК (в виде праймера, гибридизованного с матричной цепью) и приходящим для добавления к праймеру нуклеотидом. ДНК пребывает в дуплексе классической B-формы, кроме последних двух нуклеотидных пар, которые принимают более открытую A-форму. Резкий изгиб ДНК экспонирует очередное основание матрицы приходящему нуклеотиду. 3′-конец праймера (растущий) удерживается «пальцами» и «ладонью». Положение ДНК фиксируется контактами фермента с фосфодиэфирным остовом ДНК (таким образом фермент приспособлен к работе с любой нуклеотидной последовательностью).

В структурах ДНК-полимераз этого семейства при связывании одной только ДНК (без приходящего нуклеотида) «большой палец» и «пальцы» находятся в ориентации, открывающей «ладонь», с O-спиралью (O, O1, O2; см. рис. 18.8), повернутой от «ладони». Вероятно, поворот O-спирали внутрь приводит к захвату приходя-

щего нуклеотида и образованию активного каталитического центра. Когда нуклеотид связан ферментом, «пальцы» совершают 60°-поворот к «ладони», при этом концы «пальцев» смещаются на 30 Å. «Большой палец» также поворачивается к ладони (на 8°). После включения нуклеотида в ДНК обратные повороты доменов позволяют ДНК переместиться через фермент, и цикл возобновляется.

Экзонуклеазная активность ответственна за устранение неверно спаренных оснований. Активный центр экзонуклеазного домена находится далеко от «синтезирующего» активного центра. Чередование синтеза и редактирования в исполнении одного и того же фермента обусловлено конкуренцией за 3′-затравочный конец ДНК между двумя его активными центрами. Взаимодействия аминокислот активного каталитического центра с основанием неподходящего нуклеотида нарушают рабочую структуру фермента. В каталитическом центре, заблокированном неверно спаренными основаниями, «пальцы» не могут повернуться к «ладони», чтобы схватить новый приходящий нуклеотид. Это дает 3′-концу ДНК свободу связаться с экзонуклеазным доменом, что осуществляется поворотом ДНК в составе структуры фермента.

18.6 Синтез ДНК является полунепрерывным

Ключевые понятия:

- Осуществляющая репликацию ДНК-репликаза движется непрерывно, синтезируя ведущую цепь, однако синтез отстающей цепи идет посредством создания коротких фрагментов, которые затем соединяются.

Антипараллельность цепей дуплекса ДНК ставит перед репликацией следующую проблему: по мере продвижения репликативной вилки дочерние цепи должны синтезироваться на обеих экспонируемых в качестве матриц цепях дуплекса. Вилка движется в направлении 5′–3′ по одной цепи и в направлении 3′–5′ по другой цепи. Однако синтез нуклеиновой кислоты в природе всегда идет от 5′-конца к 3′-концу. Решение этой проблемы состоит в том, что цепь, не подходящая для прямой репликации, растет как целое в направлении 3′–5′, но делает это сериями коротких фрагментов, каждый из которых синтезируется «задом наперед», причем растущим, как и положено, оказывается 3′-конец.

Рассмотрим подробнее регион, расположенный непосредственно за репликативной вилкой (РИС. 18.9). Мы опишем происходящие события с учетом отличительных особенностей каждой из новосинтезированных цепей ДНК:

- Синтез **ведущей** (лидирующей)-**цепи** (leading strand) может представлять собой непрерывный «пробег» в 5′–3′-направлении, по мере раскручивания матричного дуплекса.
- Для синтеза **отстающей цепи** (lagging strand) необходим открытый отрезок одноцепочечной матрицы, на котором фрагмент ДНК растет в направлении, обратном движению вилки. После синтеза серии таких 5′–3′-фрагментов они соединяются с образованием непрерывной отстающей цепи.

Прерывистая репликация может быть отслежена с помощью короткого импульса мечения радиоактивностью. Метка обнаруживается в составе коротких фрагментов ДНК, седиментирующих в интервале от 7S до 11S, что соответствует длине в 1000–2000 оснований. Такие **фрагменты Оказаки** обнаруживаются в реплицирующейся ДНК и про-, и эукариот. После увеличения продолжительности инкубации с радиоактивностью метка входит в состав больших кусков ДНК. Это изменение связано с образованием ковалентных связей между фрагментами Оказаки.

Отстающая цепь *должна* быть синтезирована в виде фрагментов Оказаки. В течение долгого времени оставалось неясным, синтезируется ведущая цепь таким же образом или — непрерывно. Вся новосинтезированная ДНК *E. coli* обнаруживалась в виде коротких фрагментов. На первый взгляд, это значило, что обе цепи синтезируются прерывисто. Однако оказалось, что не все эти куски ДНК представляли собой *bona fide* фрагменты Оказаки; некоторые псевдофрагменты получались в результате разрывов цепи ДНК, на самом деле являвшейся продуктом непрерывного синтеза. Причиной разрывов было встраивание в ДНК некоторого количества урацила вместо тимина. После удаления урацила системой репа-

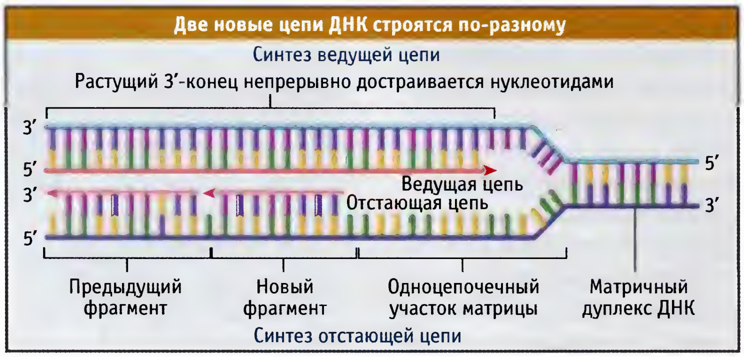


РИС. 18.9 Ведущая цепь синтезируется непрерывно, отстающая — в виде фрагментов

рации и пока на его место не встроен тимин, ведущая цепь оказывается разорванной.

Таким образом, отстающая цепь синтезируется прерывисто, а ведущая цепь — непрерывно. Это явление называют **полунепрерывной репликацией** (semidiscontinuous replication).

18.7 Наработка одноцепочечной ДНК для размножения фага в рамках модельной системы φX

Ключевые понятия:

- Для репликации необходима хеликаза — чтобы разделить цепи ДНК за счет энергии гидролиза АТФ.
- Белок SSB требуется для поддержания разделенных цепей ДНК в одноцепочечном состоянии.
- Комбинация хеликазы, SSB и белка А разделяет φX174-дуплекс на одноцепочечное кольцо и одноцепочечный линейный фрагмент.

По мере своего продвижения репликативная вилка раскручивает двойную спираль ДНК. Одна из цепей матрицы вскоре опять переходит в форму дуплекса, на этот раз образуя его с ведущей дочерней цепью, по мере ее синтеза. Другая половина матрицы дольше остается одноцепочечной, до тех пор пока достаточно протяженный ее участок не окажется открытым для синтеза фрагмента Оказаки отстающей цепи в обратном направлении. Генерация одноцепочечного состояния ДНК и его поддержание является критическим аспектом всей репликации. Для перевода двухцепочечной ДНК в одноцепочечное состояние необходимы два типа ферментативных активностей:

- **Хеликаза** — это фермент, разделяющий цепи ДНК, обычно используя гидролиз АТФ для обеспечения необходимой энергии.
- **Белок SSB** (single-strand binding protein) связывающий одноцепочечную ДНК, предотвращает переход разделенных цепей в состояние дуплекса. Молекулы SSB связываются с ДНК как мономеры, но обычно в режиме кооперации: первичный комплекс облегчает связывание последующих мономеров.

Существует множество ситуаций, в которых хеликазам приходится разделять цепи дуплекса нуклеиновой кислоты: разделение цепей в точке роста репликативной вилки, катализ миграции (рекомбинационных) структур Холлидея вдоль ДНК, и т.д. У *E. coli* есть двенадцать разных хеликаз. В общем случае хеликаза — мультимер. Самая распространенная форма хеликазы — гексамер. Он обычно перемещается вдоль ДНК, пользуясь своей мультимерной структурой для связывания с ДНК по многим точкам.

На рис. [illegible] показана обобщенная схематичная модель действия гексамерной хеликазы. Она склонна принимать одну конформацию, которая связывает дуплекс ДНК, и другую, которая связывает одноцепочечную ДНК. Их чередование приводит в действие мотор, расплавляющий дуплекс и требующий гидролиза АТФ (одной молекулы АТФ достаточно, чтобы разомкнуть одну пару оснований). Хеликаза начинает расплетать ДНК на одноцепочечном участке, примыкающем к дуплексу. Фермент может функционировать с определенной полярностью, предпочитая одноцепочечную ДНК с 3′-концом (3′–5′-хеликаза) либо с 5′-концом (5′–3′-хеликаза).

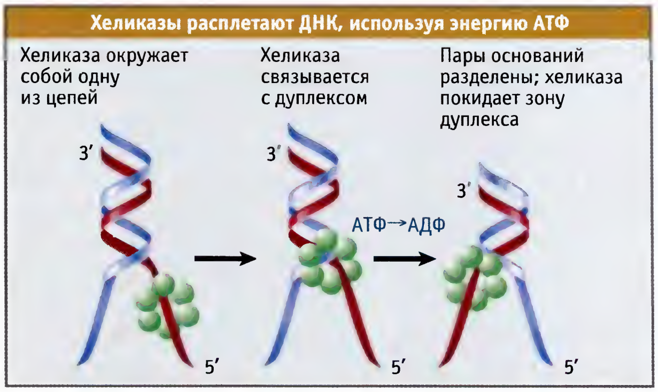


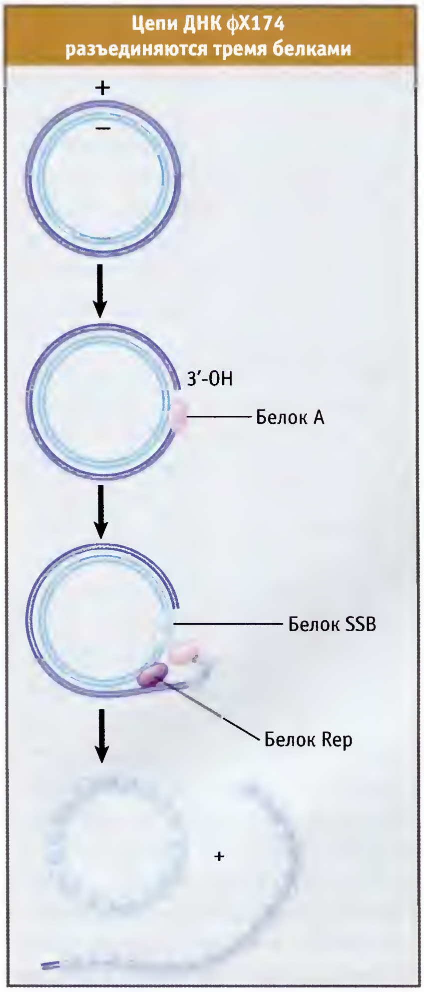
РИС. 18.10 Гексамер ДНК-хеликазы движется вдоль одной из цепей ДНК. Связываясь с двухцепочечным участком, хеликаза, скорее всего, меняет свою конформацию. Используя гидролиз АТФ, фермент разделяет цепи дуплекса, после чего возвращается в конформацию, необходимую для связывания одиночной цепи

Особенности процесса разделения цепей можно проиллюстрировать на примере хорошо изученной конверсии двухцепочечной ДНК фага φX174 в отдельные одиночные цепи. На рис. [illegible] показано, как одиночная цепь сходит с кольцевой цепи (напоминая «катящееся кольцо»), что уже было показано на рис. 16.6. Реакция может иметь место и в отсутствие репликации, достаточно доступности ДНК для трех упомянутых белков *in vitro*.

А-белок фага вносит одноцепочечный разрыв в (+)-цепь вирусной ДНК в ориджине репликации. В присутствии двух белков бактерии-хозяина (Rep и SSB) и АТФ надрезанная ДНК расплетается. Белок Rep представляет собой хеликазу, которая разделяет цепи; SSB удерживает их в одноцепочечной форме. Белок SSB *E.coli* представляет собой тетрамер в 74 кДа, который кооперативно связывается с одноцепочечной ДНК.

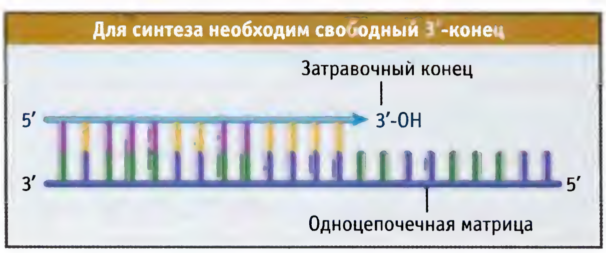
Значение кооперативности в том, что если одна молекула белка связалась, то следующей связаться уже гораздо легче. Таким образом, как только реакция связывания стартовала на данной молекуле ДНК, она стремительно распространяется, пока вся одноцепочечная ДНК не покроется белком SSB. Обратите внимание, что этот белок не является ДНК-расплетающим; его функция заключается лишь в стабилизации ДНК, уже пребывающей в одноцепочечном состоянии.

В нормальных обстоятельствах *in vivo* реакции расплетания, покрывания и репликации идут плотным строем. По мере продвижения репликативной вилки ДНК связывается SSB, который держит цепи разделенными и готовыми служить матрицами. В репликативной вилке SSB расходуется стехиометрически. Он необходим



ДНК фX174 может быть разделена на одиночные цепи комбинацией трех энзиматических активностей: белок A надрезает ДНК, хеликаза Rep расплетает ДНК, белок SSB стабилизирует одиночные цепи

для нескольких стадий репликации; *ssb*-мутанты имеют фенотип «быстрого реагирования» и проявляют нарушения не только репликации, но также рекомбинации и репарации. Некоторые фаги (например, T4) используют разные SSB; это указывает на специфические взаимодействия между SSB и компонентами репликационного аппарата (см. разд. 18.14 «Фаг T4 обладает собственным репликативным аппаратом»).



Чтобы начать репликацию, ДНК-полимеразе необходим 3′-OH конец

18.8 Чтобы начать синтез ДНК, необходима затравка

Ключевые понятия:

- Всем ДНК-полимеразам требуется 3′-OH затравочный конец.
- Затравочный конец может быть обеспечен РНК-затравкой (праймером), одноцепочечным разрывом в ДНК или затравляющим белком.
- Для репликации ДНК специальная РНК-полимераза (праймаза) синтезирует цепь РНК, обеспечивающую затравочный конец.
- У *E. coli* начало синтеза ДНК может происходить двумя способами: в бактериальном ориджине (oriC) и в ориджине фX174.
- Затравление репликации на двухцепочечной ДНК всегда требует наличия репликазы, SSB и праймазы.
- DnaB — хеликаза, разворачивающая ДНК при репликации *E. coli*

Ни одна ДНК-полимераза «не умеет» инициировать синтез цепи *de novo*. На рис. 18.12 изображены особенности процесса инициации. Синтез новой цепи может быть начат лишь со свободного 3′-OH конца, «заглубленного» по отношению к комплементарной (матричной)-цепи, которая должна быть переведена в одиночное состояние.

Участок цепи со свободным 3′-OH концом называется **затравкой**, или праймером. Он может «принимать» различные формы. Типы затравочных реакций представлены на рис. :

- На матрице синтезируется последовательность РНК, а потом ДНК-полимераза начинает удлинять ее 3′-OH-конец. Этот способ в основном и используется для репликации клеточной ДНК, а также ДНК некоторых вирусов (см. рис. 17.18 в разд. 17.11 «Система совместимости ColE1 контролируется РНК-регулятором»).
- Уже готовая ДНК спаривается с матрицей, позволяя своему 3′-OH-концу послужить затравкой для синтеза ДНК. Этот механизм используется ретровирусами для затравления обратной транскрипции РНК (см. рис. 22.6 в разд. 22.4 «ДНК ретровируса является продуктом обратной транскрипции»).
- Затравочный конец образуется посреди ДНК-дуплекса. Самый распространенный способ — внесение ника, как для инициации «катящегося кольца» (см. рис. 16.5). В этом случае одна из цепей дуплекса по мере нового синтеза оказывается вытесненной. (Обратите внимание на отличие от ник-трансляции (рис. 18.5), в ходе которой находящаяся в нике (одноцепочечном разрыве) ДНК-полимераза I одновременно занимается и синтезом, и деградацией ДНК.)
- Реакцию затравляет белок, «навязывающий» нуклеотид ДНК-полимеразе. Эту реакцию используют некоторые вирусы (см. рис. 17.18 в разд. 17.11 «Система совместимости ColE1 контролируется РНК-регулятором»).

Этап затравления, приводящий к образованию 3′-OH-конца, требуется для запуска синтеза как ведущей, так и отстающей цепей. Ведущая цепь нуждается в разовой инициации, которая имеет место в ориджине. Для отстающей цепи необходима серия инициаций, т. к. для каждого фрагмента Оказаки запуск синтеза идет *de novo*. Каждый фрагмент Оказаки начинается последовательностью

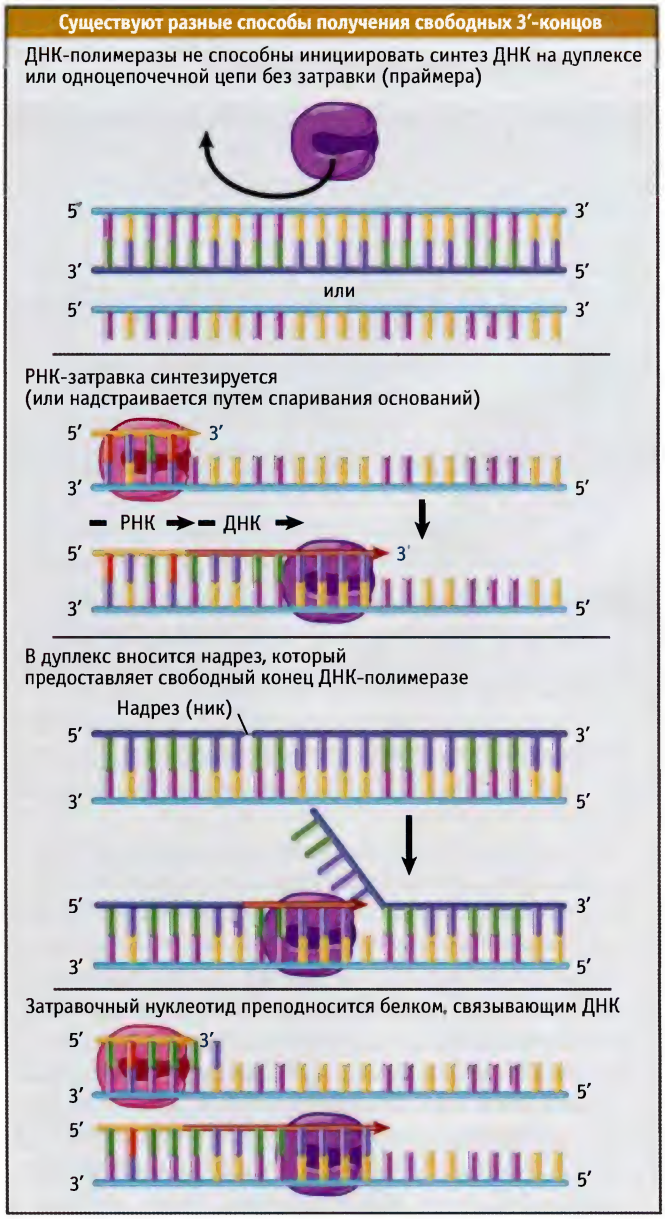


РИС. 18.13 **Для начала синтеза ДНК-полимеразе необходим затравочный 3′-конец; его можно получить несколькими способами**

РНК-затравки, размером примерно с десяток оснований. ДНК-полимераза наращивает ее 3′-OH-конец.

Над образованием затравки работает фермент **праймаза**. У *E.coli* это продукт гена *dnaG*, он обладает особой РНК-полимеразной активностью и представляет собой одиночный полипептид в 60 кДа (гораздо мельче РНК-полимеразы). Праймаза ипользуется для полимеризации рибонуклеотидов лишь в особых обстоятельствах, а именно для синтеза коротких фрагментов РНК, используемых в качестве праймеров для синтеза ДНК. Праймаза DnaG кратковременно ассоциирует с репликативным комплексом и синтезирует праймер (обычно из 11–12 оснований). Он начинается последовательностью pppAG, напротив триплета 3′-GTC-5′ матрицы.

Некоторые системы пользуются альтернативами праймазе DnaG. Так, между двумя фагами, M13 и G4, служившими объектами ранних исследований по репликации, было обнаружено интересное различие. У G4 для затравления используется DnaG, а у M13 — бактериальная РНК-полимераза. Оба фага имеют характерную особенность: сайт прайминга отмечен особой вторичной структурой ДНК.

У *E. coli* обнаружено два типа затравления:

- Система *oriC* (по названию бактериального сайта начала репликации) основана на ассоциации праймазы DnaG с белковым комплексом в репликативной вилке.
- Система φ*X* (по названию фага φX174) требует наличия комплекса инициации, состоящего из дополнительных компонентов — так называемой праймосомы (см. разд. 18.17 «Праймосома перезапускает репликацию»).

Поэтому иногда о репликонах говорят, что они φ*X* либо *oriC* типа.

Типы ферментативной активности, принимающей участие в инициации, обобщает РИС. 18.14. Хотя в некоторых репликонах *E. coli* конкретные белки могут заменяться альтернативными, одни и те же типы активности необходимы всегда. Хеликаза требуется для образования одиночных цепей, белок SSB поддерживает цепи в одиночном состоянии, праймаза синтезирует РНК-затравку.

DnaB является центральным компонентом и в φ*X*, и в *oriC* репликонах. Этот белок обладает 5′–3′-хеликазной активностью, расплетающей ДНК. Энергоснабжение реакции происходит за счет расщепления АТФ. DnaB, по сути, представляет собой активный компонент «точки роста». В репликонах *oriC* DnaB исходно садится на ориджин в составе большого комплекса (см. разд. 18.15 «Формирование репликативных вилок в ориджине»). DnaB образует точку роста, в которой цепи матричного дуплекса разделяются по мере продвижения репликативой вилки. Он является частью ДНК-полимеразного комплекса и взаимодействует с DnaG-праймазой, инициируя синтез фрагментов Оказаки отстающей цепи.

18.9 Холофермент ДНК-полимеразы состоит из трех субкомплексов

Ключевые понятия:

- **Репликаза *E. coli* ДНК-полимераза III — это 900 кДа комплекс с димерной структурой.**
- **Каждая мономерная субъединица обладает каталитическим центром, димеризационной субъединицей и компонентом процессивности.**
- **Специальный белок загружает субъединицы процессивности на ДНК, и они зажимают ее в кольцо.**
- **На каждую цепь матрицы приходится по одному каталитическому центру.**

Теперь мы можем соотнести субъединичную структуру ДНК-полимеразы III *E. coli* со всеми необходимыми для синтеза ДНК активностями (в свете их недавнего рас-

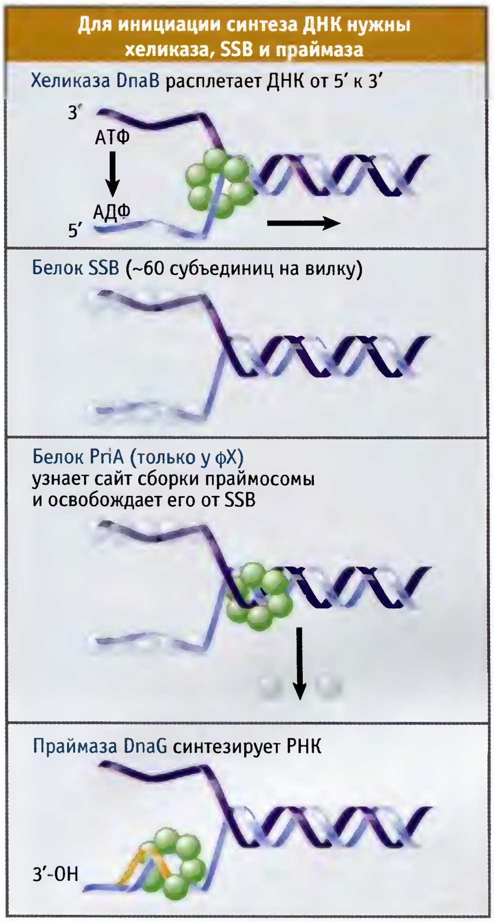


РИС. 18.14 Инициация репликации обеспечивается совместными усилиями целого ряда различных белков

смотрения) и попытаться смоделировать ее действие. Холофермент (реплисома) представляет собой 900 кДа комплекс, содержащий десять белков, подразделенных на четыре типа субкомплексов:

- Две копии каталитической «сердцевины». Каждая «сердцевина» содержит субъединицу α (с ДНК-полимеразной активностью), субъединицу ε (3′–5′ корректирующую экзонуклеазу) и субъединицу θ (стимулирующую экзонуклеазу).
- Две копии димеризующей субъединицы τ, которые служат для скрепления минимального фермента.
- Две копии *зажима* (clamp), которые ответственны за удержание минимального фермента на их ДНК-матрицах. Каждый зажим состоит из гомодимера субъединиц β, замкнутого в обхват ДНК, обеспечивая процессивность.
- Комплекс γ представляет собой группу из пяти белков, образующих **клэмп-лоудер** — приспособление для накладывания зажима на ДНК.

Модель сборки ДНК-полимеразы III изображена на РИС. 18.15. Холофермент монтируется на ДНК в три этапа:

- Сначала клэмп-лоудер соединяет β-субъединицы с комплексом праймер–матрица, используя энергию гидролиза АТФ.
- Связывание с ДНК меняет конформацию β-субъединицы (в том месте, где она контактирует с клэмп-лоудером) так, что она обнаруживает высокое сродство к минимальному ферменту (полимеразе). В результате полимераза тоже оказывается прикрепленной к ДНК.
- Димер τ связывается с ферментом (полимеразой) и выполняет свою димеризующую функцию, связывая второй минимальный фермент (полимеразу), ассоциированную со вторым β-зажимом. Клэмп-лоудер отвечает за окружение каждой из двух цепей матрицы парой β-димеров, а также за снятие β-комплекса с ДНК по завершении синтеза.

Холофермент ДНК-полимеразы III асимметричен, поскольку клэмп-лоудер в нем всего один. Вследствие этой асимметрии минимальный фермент в его половинках различаются по своей способности «расставаться» с ДНК. Это соответствует различию стоящих перед ними задач: одна полимераза синтезирует непрерывную ведущую цепь (не «отпуская» матрицу), а другая — фрагментарную отстающую цепь (то и дело диссоциируя и реассоциируя с матрицей). Клэмп-лоудер находится при той субъединице, которая строит отстающую цепь, оказывая ей ключевое содействие в синтезе обособленных фрагментов Оказаки.

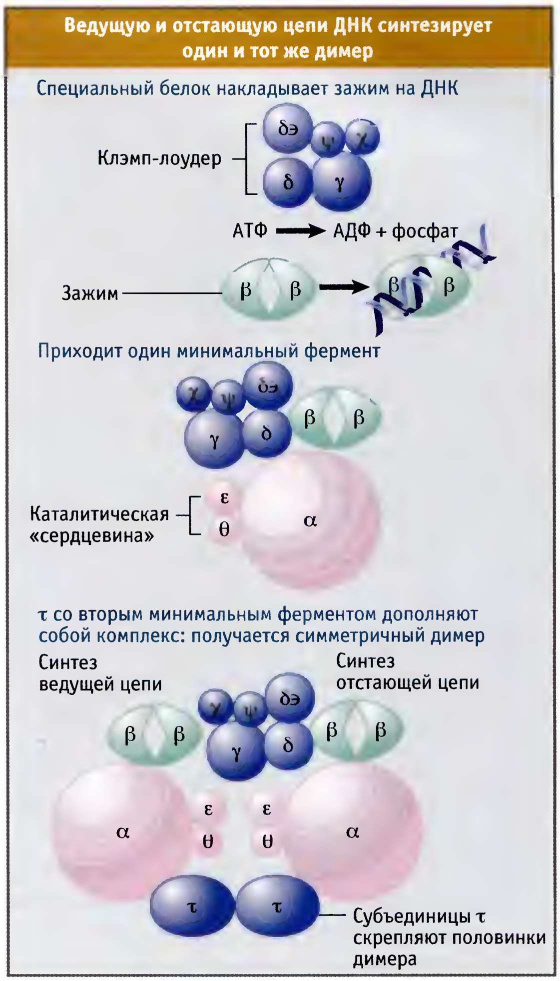


РИС. 18.15 Сборка холофермента ДНК-полимеразы III проходит в несколько стадий. Холофермент синтезирует обе новые цепи ДНК

18.10 Зажим контролирует ассоциацию минимального фермента с ДНК

Ключевые понятия:

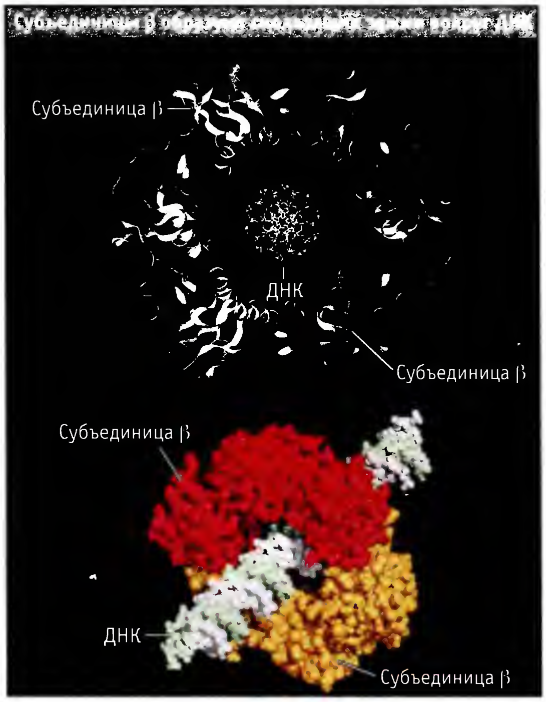
- Полимераза (минимальный фермент) ведущей цепи отличается высокой процессивностью, поскольку зажим удерживает ее на ДНК.
- Зажим, ассоциированный с полимеразой отстающей цепи, диссоциирует в конце каждого фрагмента Оказаки и заново собирается для следующего фрагмента.
- Хеликаза DnaB отвечает за взаимодействие с праймазой DnaG для инициации синтеза фрагмента Оказаки.

Благодаря β-субъединицам холофермент ДНК-полимеразы III отличается высокой процессивностью. β-субъединицы образуют кольцеобразный димер, крепко связанный с ДНК, но способный скользить вдоль ее дуплекса. На представлена модель β-кольца по отношению к двойной спирали ДНК, созданная на основании данных кристаллографического анализа. Кольцо имеет внешний диаметр 80 Å и внутреннюю полость диаметром 35 Å — почти вдвое шире двойной спирали ДНК (20 Å). Пространство между кольцом белка и ДНК заполнено водой. Каждая β-субъединица имеет три глобулярных домена со сходной организацией (хотя их аминокислотные последовательности отличаются). Вследствие этого димер имеет шестикратную симметрию, отображенную двенадцатью α-спиралями, выстилающими внутреннюю поверхность кольца.

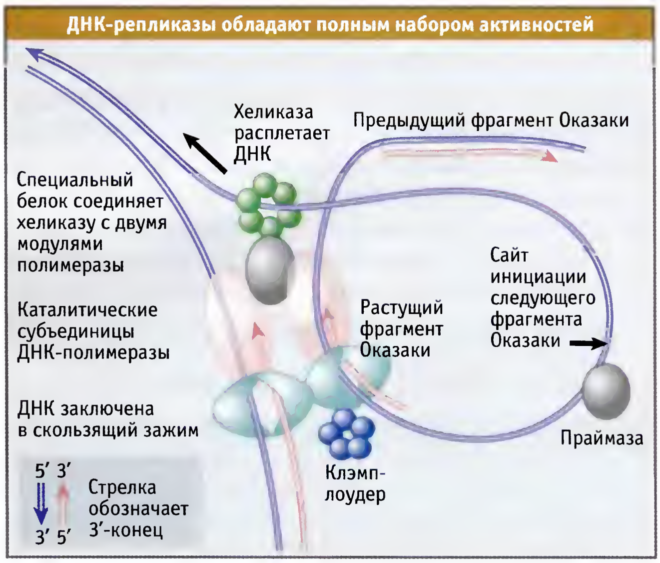
Димер окружает дуплекс, образуя скользящий зажим, позволяющий холоферменту скользить вдоль ДНК. Такая структура объясняет высокую процессивность — фермент просто не может соскочить с ДНК! Положительные заряды в α-спиралях на внутренней стороне кольца могут взаимодействовать с ДНК через посредничество молекул воды (которое облегчает установление и разрыв контактов). Белковый зажим не контактирует с ДНК напрямую, так что может ехать вдоль нее. как на салазках. с молекулами воды в качестве смазки.

Как зажим накладывается на ДНК? Он представляет собой кольцо субъединиц, обхватывающих ДНК; его сборка или удаление энергозависимы, что обеспечивается клэмп-лоудером (комплекс γ). Он представляет собой замкнутый пентамер, связывающий открытую форму β-кольца, готовую для захвата ДНК. δ-субъединица клэмп-лоудера открывает кольцо по одному из стыков между β-субъединицами. Таким образом, клэмп-лоудер садится на верхушку закрытого зажима, обращая к нему свой АТФазный активный центр, и расходует энергию гидролиза АТФ на открывание зажима и вкладывание ДНК в его центральную полость.

Базовый принцип модели двойственного функционирования репликазы состоит в том, что в то время как одна ее половинка непрерывно занята синтезом ведущей цепи, другая в циклическом режиме инициирует и терминирует фрагменты Оказаки отстающей цепи внутри большой одноцепочечной петли, которую образует ее половинка матрицы. На приведена



Димер субъединиц β холофермента ДНК-полимеразы III охватывает дуплекс ДНК замкнутым кольцом. На схеме ДНК находится в центре, две субъединицы β, расположенные в ориентации «голова-хвост», показаны оранжевым и красным цветами. Перепечатано из Cell, vol. 69, Kong, X. P., et al., pp. 425–437. Copyright 1992, с разрешения Elsevier. Любезно предоставлено John Kuriyan, University of California, Berkeley



Хеликаза, адаптирующая ДНК под репликативную вилку, соединена с двумя каталитическими субъединицами ДНК-полимеразы. Каждая из них удерживается на ДНК скользящим зажимом. Полимераза, которая синтезирует ведущую цепь, движется по ДНК непрерывно. Полимераза, которая синтезирует отстающую цепь, сходит с ДНК в конце каждого фрагмента Оказаки, чтобы снова ассоциировать с праймером в петле, образуемой одиночной цепью матрицы, для синтеза следующего фрагмента Оказаки

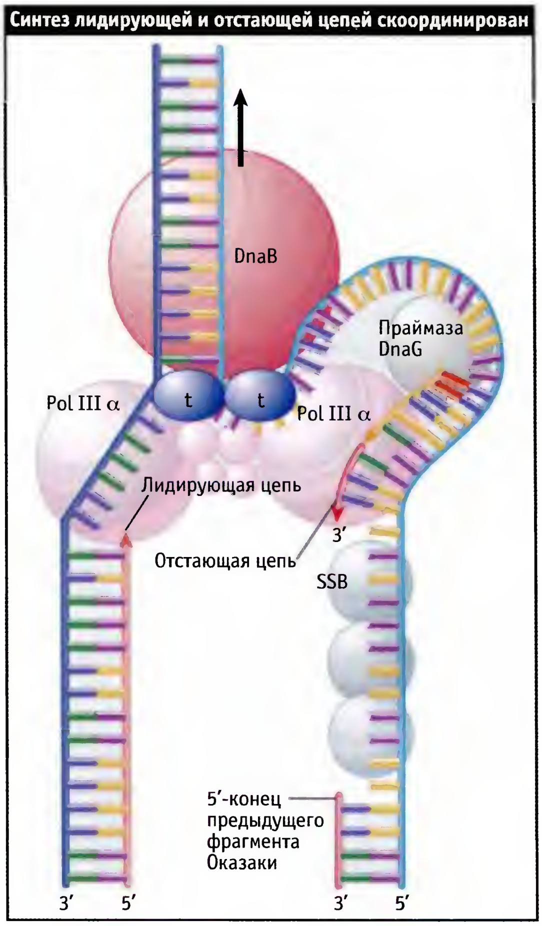


РИС. 18.18 Каждый минимальный фермент Pol III синтезирует дочернюю цепь ДНК. Белок DnaB отвечает за продвижение всей репликативной вилки

обобщенная схема работы такой репликазы. Репликативную вилку создает хеликаза (в характерной форме — гексамер-кольцо), которая перемещается в направлении 5′–3′ по матрице отстающей цепи. Хеликаза соединяется с двумя каталитическими субъединицами ДНК-полимеразы, каждая из которых, в свою очередь, ассоциирована со скользящим зажимом.

Для ДНК-полимеразы III эта модель может быть описана в терминах индивидуальных компонентов ферментного комплекса, это показано на РИС. 18.18. С каждой цепью матрицы ассоциировано по одному минимальному ферменту. Холофермент непрерывно движется вдоль матрицы для ведущей цепи; матрица для отстающей цепи «протаскивается» с отставанием, образуя петлю. DnaB создает точку расплетания, перемещаясь вниз по ходу ДНК.

DnaB контактирует с τ-субъединицей (τ-субъединицами) клэмп-лоудера, устанавливая таким образом прямую связь между комплексом хеликазы-праймазы и минимальными ферментами ДНК-полимеразы III. У такого сцепления есть два эффекта. Первый — увеличение скорости движения минимального фермента в десять раз. Второй — предотвращение соскакивания полимеразы с ведущей цепи, т. е. повышение процессивности.

Синтез ведущей цепи приводит к образованию петли одноцепочечной ДНК, служащей матрицей для синтеза отстающей цепи, и эта петля увеличивается по мере продвижения точки расплетания. После инициации фрагмента Оказаки полимераза отстающей цепи протягивает одноцепочечную матрицу сквозь β-зажим, синтезируя новую цепь. Таким образом, петля одноцепочечной матрицы вмещает по меньшей мере один фрагмент Оказаки, вплоть до окончания его синтеза и приведения полимеразы в готовность начать следующий.

Что происходит по завершении фрагмента Оказаки? Все компоненты репликативного аппарата функционируют процессивно (т. е. пребывают ассоциированными с ДНК), за исключением праймазы и β-зажима. На РИС. 18.9 показано, как они диссоциируют по завершении синтеза фрагмента, отпуская петлю. Вскоре клэмплоудер привлекает новый β-зажим для инициации следующего фрагмента Оказаки. В каждом таком цикле полимераза отстающей цепи переходит из одного β-зажима в другой, не покидая репликативный комплекс.

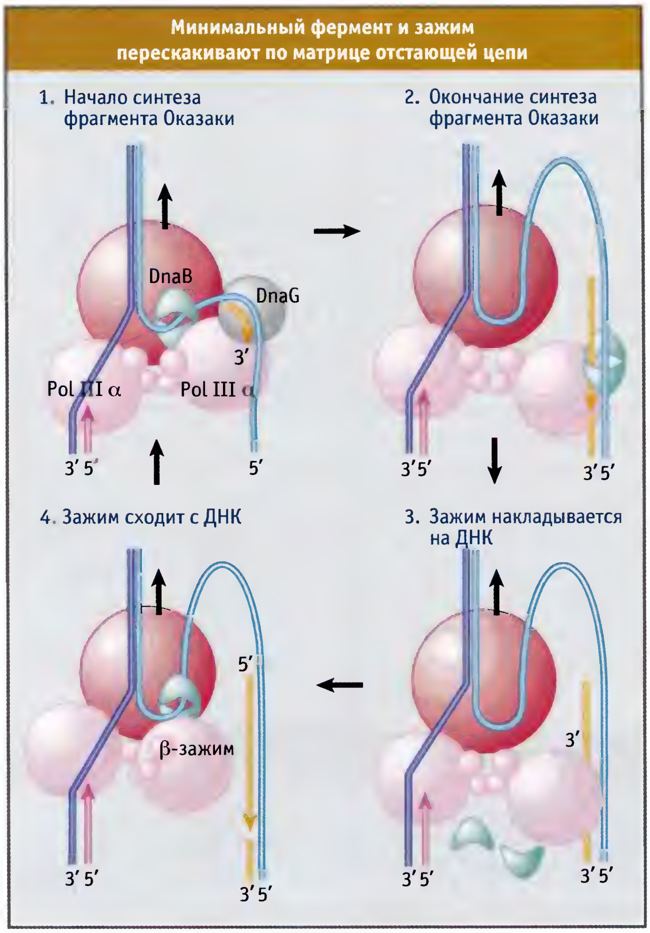


РИС. 18.19 По завершении синтеза очередного фрагмента Оказаки скользящий зажим отпускает отстающую цепь с матрицей. Тот же зажим вместе со своим минимальным ферментом накладывается на ДНК в сайте инициации синтеза следующего фрагмента Оказаки

Что отвечает за распознавание сайтов инициации синтеза фрагментов Оказаки? В репликонах *oriC* связь прайминга с репликативной вилкой обеспечивается двоякими свойствами DnaB: будучи хеликазой, продвигающей репликативную вилку, он в то же время производит взаимодействие с праймазой. После синтеза затравки праймаза освобождается; с ДНК ее смещает, по-видимому, сама ДНК-полимераза III. Длина затравочной РНК колеблется в пределах от восьми до четырнадцати оснований.

18.11 Координирование синтеза ведущей и отстающей цепей

Ключевые понятия:

- Для синтеза ведущей и отстающей цепей требуются разные комплексы ферментов.
- У *E. coli* оба этих комплекса содержат одну и ту же каталитическую субъединицу (DnaE).
- У других организмов для каждой из цепей могут требоваться различные каталитические субъединицы.

Каждая новая цепь ДНК синтезируется индивидуальным каталитическим модулем. РИС. 18.20 показывает, что эти два модуля ведут себя по-разному, поскольку новые цепи ДНК растут в противоположных направлениях. Один ферментный модуль движется вместе с точкой расплетания и синтезирует ведущую цепь непрерывно. Другой модуль движется назад относительно ДНК, вдоль открытой одиночной цепи. В каждый отдельный момент времени открыт лишь небольшой кусок матрицы. По завершении синтеза одного фрагмента Оказаки необходимо начать синтез следующего где-то неподалеку от точки роста ведущей цепи. Это, в свою очередь, требует транслокации ферментного модуля, синтезирующего отстающую цепь, относительно ДНК.

Термин «ферментный модуль» оставляет открытым вопрос об однотипности ферментов, синтезирующих ведущую и отстающую цепи. У *E. coli*, о которой мы наиболее осведомлены, есть только одна разновидность каталитических субъединиц ДНК-полимеразы, используемых в репликации — белок DnaE. Активная репликаза представляет собой димер, и каждая половинка димера



РИС. 18.20 Полимеразы ведущей и отстающей цепей ДНК разнесены в пространстве

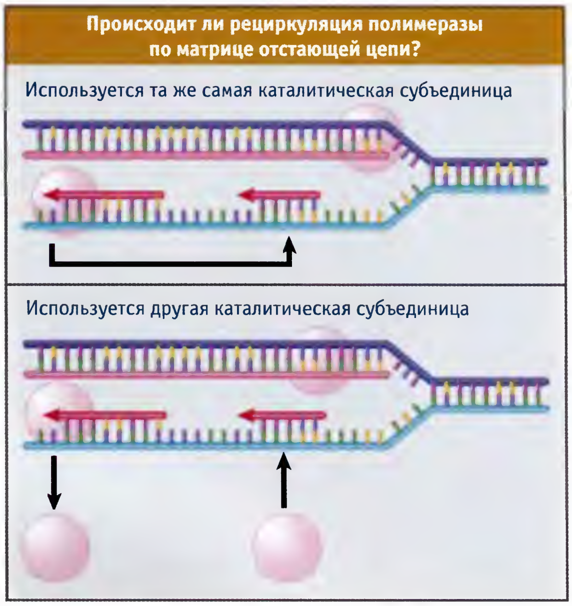


РИС. 18.21 Модель перемещений полимеразы по матрице, соответствующая верхней части рисунка, требует непрерывной рециркуляции одного и того же ферментного модуля. В соответствии с нижней частью рисунка, синтезом отстающей цепи занимается множество ферментных модулей: каждый новый фрагмент Оказаки синтезируется новым модулем

содержит DnaE в качестве каталитической субъединицы. DnaE функционирует при поддержке других белков (разных для ведущей и отстающей цепей).

Использование репликазой однотипных каталитических субъединиц, возможно, атипично. У бактерии *Bacillus subtilis* присутствуют две разных каталитических субъединицы. PolC — гомолог DnaE *E. coli* — ответственна за синтез ведущей цепи. Родственный белок, $DnaE_{BS}$, является каталитической субъединицей, синтезирующей отстающую цепь. Репликазы эукариот придерживаются той же схемы (синтез ведущей и отстающей цепей отдельными комплексами белков), однако неясно, одинаковы ли каталитические субъединицы этих комплексов (см. разд. 18.13 «Инициация и элонгация у эукариот осуществляется разными полимеразами»).

Важнейшей задачей полунепрерывной системы, использующей отдельный ферментный модуль для синтеза каждой из цепей, является бесперебойная синхронизация работы этих модулей. Иными словами, как синтез ведущей цепи оказывается скоординирован с синтезом отстающей? На РИС. 18.21 предложено две возможных модели. Согласно первой, один и тот же комплекс реутилизируется для синтеза всей череды фрагментов Оказаки. Согласно второй, комплекс диссоциирует с ДНК, чтобы другой комплекс смог присоединиться и синтезировать новый фрагмент Оказаки. Как мы видели в разд. 18.10 «Зажим контролирует ассоциацию минимального фермента с ДНК», приемлема именно первая модель.

18.12 Фрагменты Оказаки соединяются лигазой

Ключевые понятия:

- Каждый фрагмент Оказаки начинается праймером (затравкой) и заканчивается перед следующим фрагментом.
- ДНК полимераза I устраняет праймер и перемещает его — этот процесс напоминает ник-трансляцию.
- ДНК-лигаза создает связь, соединяющую 3′-конец фрагмента Оказаки с 5′-началом следующего фрагмента.

Теперь можно пополнить представление о том, как происходит соединение фрагментов Оказаки. Общий порядок событий следующий: синтез РНК-затравки (РНК-праймера), продолжение ее в виде ДНК, удаление РНК-затравки, замена ее сегментом ДНК и ковалентная сшивка соседних фрагментов Оказаки.

Как показано на РИС. 18.22, синтез фрагмента Оказаки терминируется впритык к началу РНК-праймера предыдущего фрагмента. Если праймер убрать, останется «брешь». Его заполняет (скорее, не дает бреши образоваться) ДНК-полимераза I; мутанты *pol*⁻ не способны должным образом соединять фрагменты Оказаки. 5′–3′ экзонуклеазная активность убирает РНК-праймер, одновременно заменяя его последовательностью ДНК — продолжением 3′-ОН конца следующего фрагмента Оказаки. Этот процесс эквивалентен ник-трансляции, за исключением того, что новосинтезированный материал смещает с матрицы сегмент РНК, а не ДНК.

У млекопитающих (где ДНК-полимераза лишена 5′–3′ экзонуклеазной активности) фрагменты Оказаки соединяются в два этапа. Синтез фрагмента смещает РНК-затравку предыдущего фрагмента, оставляя его «свисать». На РИС. 18.23 показано, как фермент FEN1 полностью удаляет этот «хвост». В этой реакции FEN1 работает как эндонуклеаза, но, помимо этого, он обладает 5′–3′-экзонуклеазной активностью. При репарации FEN1 может сделать надрез рядом со смещенным нуклеотидом, а затем, воспользовавшись своей экзонуклеазной активностью, очистить окрестность.

Задержка с удалением «хвоста» может привести к серьезным последствиям на участках с повторяющимися последовательностями. Прямые повторы могут оказаться смещены относительно матрицы и ошибочно спарены с

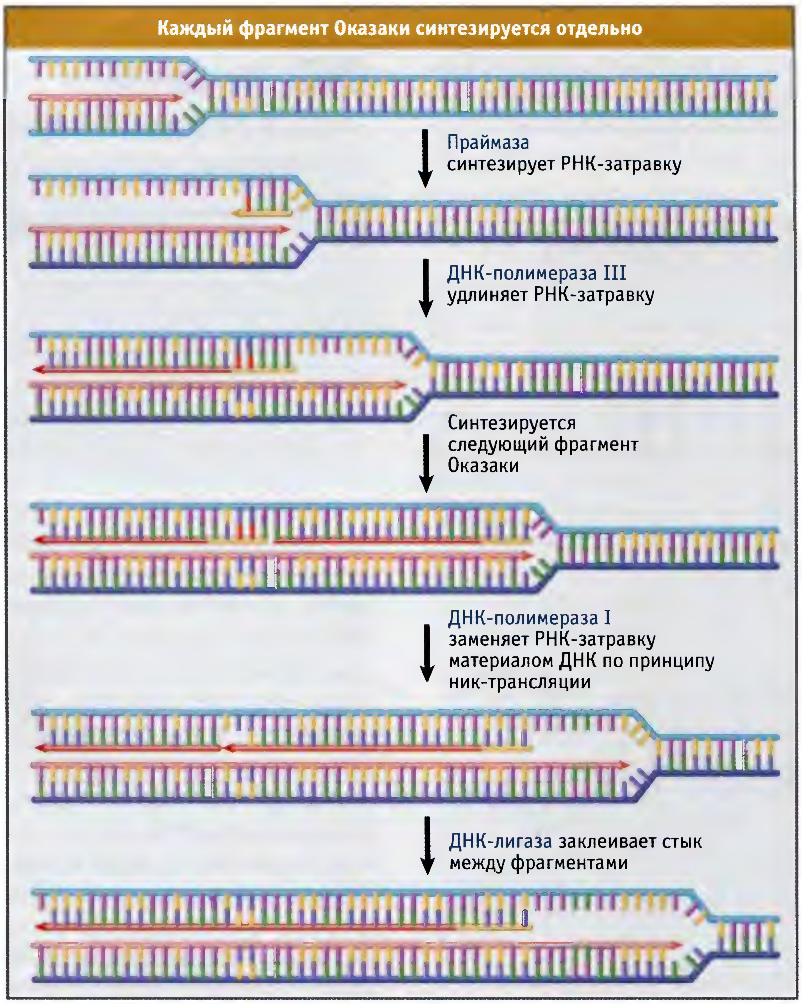


РИС. 18.22 Синтез фрагмента Оказаки включает в себя синтез РНК-затравки, ее наращивание материалом ДНК, удаление затравки предыдущего фрагмента, заклеивание стыка

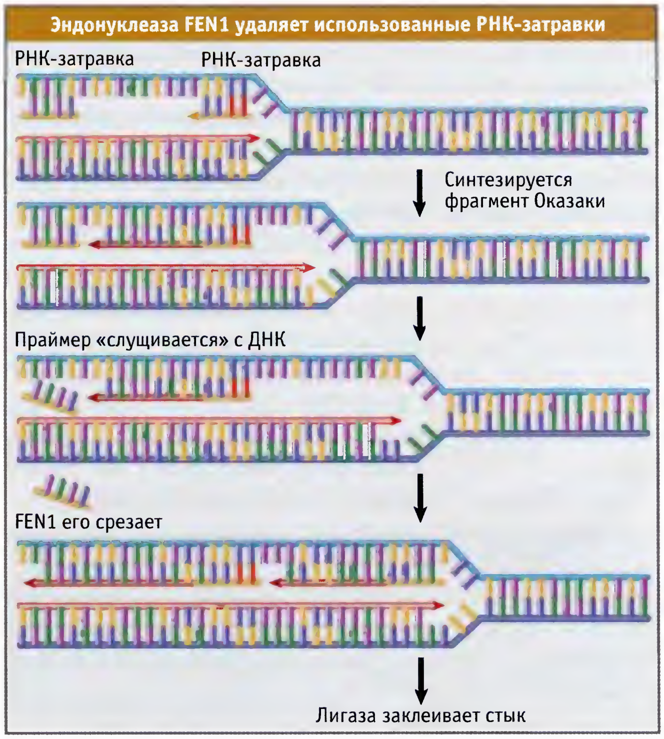


РИС. 18.23 FEN1 является нуклеазой, работающей с одиночной цепью, выбившейся из дуплекса под действием замещения новосинтезированной цепью. Срезая эти «хвосты», FEN1 освобождает новосинтезированную ДНК от использованных затравок

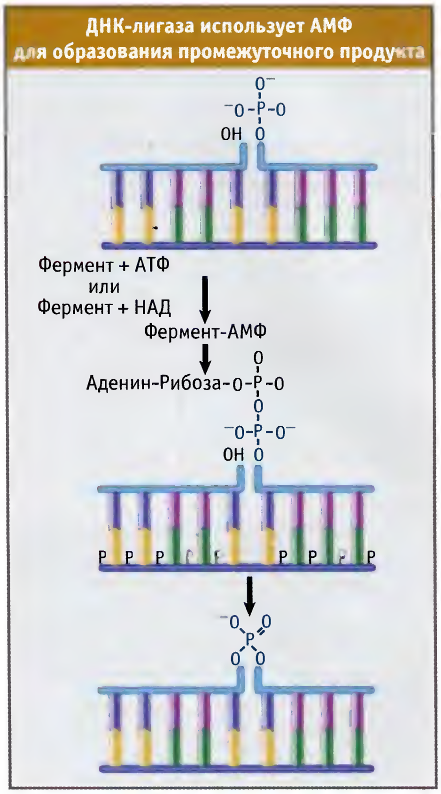


РИС. 18.24 ДНК-лигаза заклеивает стык между соседними нуклеотидами

ней на другом отрезке; палиндромные последовательности могут образовывать шпильки. Это создает угрозу изменения числа повторов после репликации (см. рис. 6.28). Удаляя «хвост», FEN1, таким образом, предотвращает индукцию появления в геноме делеций или дупликаций.

После удаления/замещения РНК-затравки фрагменты Оказаки необходимо сшить вместе. 3′-гидроксильный конец каждого фрагмента примыкает к 5′-фосфатному концу предыдущего. Этот стык зашивает фермент **ДНК-лигаза**. Лигазы есть и у про-, и у эукариот. У мутантов lig⁻ фрагменты Оказаки остаются несоединенными.

ДНК-лигазы *E. coli* и T4 обладают общей способностью сшивать стыки между 3′-гидроксильным и 5′-фосфатным концами, что показано на РИС. 18.24. Оба фермента проводят двухступенчатую реакцию, образуя комплекс с АМФ. Ферменты *E. coli* и T4 используют разные кофакторы. Фермент *E. coli* использует НАД (никотинамидадениндинуклеотид) в качестве кофактора, в то время как фермент T4 использует АТФ. АМФ в составе ферментного комплекса крепится на стыке к фосфату, между фосфатом и 3′-гидроксилом образуется фосфодиэфирная связь, что ведет к высвобождению фермента и АМФ.

18.13 Инициация и элонгация у эукариот осуществляются разными полимеразами

Ключевые понятия:

- Репликативная вилка содержит один комплекс ДНК-полимераза α/праймаза и два комплекса ДНК-полимераза δ и/или ε.
- Комплекс ДНК-полимераза α/праймаза инициирует синтез обеих цепей ДНК.
- ДНК-полимераза δ удлиняет ведущую цепь, а вторая ДНК-полимераза δ (или ДНК-полимераза ε) удлиняет отстающую цепь.

Эукариотические клетки располагают большим числом ДНК-полимераз, которые в общих чертах можно поделить на требующиеся для полуконсервативной репликации и участвующие в репарации. Репликация ядерной ДНК требует участия ДНК-полимераз α, δ и ε. Все остальные ядерные ДНК-полимеразы задействованы в синтезе сегментов новой ДНК на замену поврежденному материалу. Из таблицы на РИС. 18.25 следует, что все ядерные репликазы — это большие гетеротетрамерные ферменты. В каждом случае одна из субъединиц отвечает за катализ, остальные задействованы во вспомогательных функциях вроде прайминга, процессивности, коррекции. Все эти ферменты реплицируют ДНК с высокой точностью, то же делает и чуть менее комплексный митохондриальный фермент. Репарационные полимеразы устроены гораздо проще, многие из них состоят из одной субъединицы-мономера (которая, однако, может функционировать в контексте комплекса других ферментов репарации). Из полимераз, участвующих в репарации, только ДНК-полимераза

ДНК-полимераза	Функция	Структура
	Репликация с высокой точностью	
α	Репликация в ядре	350 кДа тетрамер
δ	"	250 кДа тетрамер
ε	"	350 кДа тетрамер
γ	Репликация в митохондриях	200 кДа димер
	Репарация с высокой точностью	
β	Репарация вырезанного нуклеотида	39 кДа мономер
	Репарация с пониженной точностью	
ζ	Прочтение тиминовых димеров	гетеромер
η	Репарация поврежденного основания	мономер
ι	Используется во время мейоза	мономер
κ	Делеция и замена оснований	мономер

ДНК-полимеразы осуществляют репликацию и репарацию

РИС. [illegible] В эукариотических клетках есть множество различных ДНК-полимераз. Работа ферментов репликации отличается высокой точностью. Ферменты репарации (за исключением ДНК-полимеразы β) допускают меньшую точность воспроизведения нуклеотидной последовательности. Ферменты репликации входят в состав обширных белковых комплексов, где каждая субъединица имеет свою конкретную активность. Ферменты репарации организованы в более простые структуры

β по точности синтеза близка к репликазам: все остальные делают ошибки гораздо чаще. Синтез митохондриальной ДНК целиком (включая репарацию и рекомбинацию, наравне с репликацией) обеспечивает ДНК-полимераза γ.

Все три ядерных ДНК-репликазы выполняют разные функции:

- ДНК-полимераза α инициирует синтез новой цепи.
- ДНК-полимераза δ элонгирует ведущую цепь.
- ДНК-полимераза ε может участвовать в синтезе отстающей цепи, однако выступает и в других ролях.

ДНК-полимераза α является особенной, поскольку способна инициировать новую цепь. Она используется для инициации обеих, ведущей и отстающей, цепей. Фермент существует в виде комплекса из 180 кДа каталитической субъединицы, ассоциированной с β-субъединицей (необходимой, по-видимому, для сборки) и двумя меньшими белками с праймазной активностью. Иногда эту полимеразу называют полимераза α/праймаза, что отражает совмещение функций затравления и удлинения цепей ДНК.

Полимераза α/праймаза связывается в ориджине с комплексом инициации и синтезирует короткую последовательность длиной около 10 рибонуклеотидов, за которыми следуют 20–30 дезоксирибонуклеотидов (иногда ее называют иДНК). Затем их замещает фермент элонгации. Для ведущей цепи это ДНК-полимераза δ. Этот момент репликации называют «переключением полимераз». Он обусловлен взаимодействиями между несколькими компонентами инициативного комплекса.

ДНК-полимераза δ — это высокопроцессивный фермент, который в непрерывном режиме синтезирует ведущую цепь. Такая высокая процессивность — результат взаимодействия с двумя другими белками, RF-C и PCNA.

Функции RF-C и PCNA аналогичны функциям γ и β компонетов холофермента ДНК-полимеразы III *E. coli* (см. разд. 18.10 «Зажим контролирует ассоциацию минимального фермента с ДНК»). Белок RF-C катализирует наложение PCNA (зажима) на ДНК. Он связывается с 3′-концом иДНК и, пользуясь гидролизом АТФ, открывает кольцо PCNA для «обхвата» ДНК. Процессивность ДНК-полимеразы δ поддерживается белком PCNA, который прикрепляет этот фермент к матрице. (PCNA, proliferating cell nuclear antigen — ядерный антиген пролиферирующих клеток.) По своей кристаллической структуре PCNA напоминает β-зажим *E. coli*: тример образует кольцо вокруг ДНК. При всем отличии от димерного β-зажима по аминокислотной последовательности и субъединичной организации, PCNA выполняет ту же функцию.

С отстающей цепью все не так определенно. Возможно, элонгацию отстающей цепи проводит та же ДНК-полимераза δ. Она способна димеризоваться и способна быть аналогом репликазы *E. coli* (см. разд. 18.9, «Холофермент ДНК-полимеразы состоит из трех субкомплексов»). Некоторые данные указывают на то, что отстающую цепь может элонгировать ДНК-полимераза ε, хотя ей также приписывают и другие роли.

Общая модель подразумевает репликативную вилку с одним комплексом ДНК-полимеразы/праймазы и двумя другими ДНК-полимеразными комплексами. Одна из них — ДНК-полимераза δ, другая — или δ, или ε. Два комплекса ДНК-полимеразы δ/ε ведут себя так же, как два комплекса ДНК-полимеразы III в реплисоме *E. coli*: одна синтезирует ведущую цепь, другая — фрагменты Оказаки отстающей цепи. Экзонуклеаза MF1 удаляет РНК-затравки фрагментов Оказаки. ДНК-лигаза I сшивает стыки между готовыми фрагментами.

18.14 Фаг T4 обладает собственным репликативным аппаратом

Ключевые понятия:

- Фаг T4 обладает собственным репликативным аппаратом, состоящим из ДНК-полимеразы, p32 SSB, хеликазы, праймазы и сопутствующих белков, увеличивающих скорость и процессивность.

В результате внедрения фага T4 в клетку *E. coli*, несколько привнесенных им функциональных систем замещают или затрудняют соответствующие клеточные функции бактерии-хозяина. Функции хозяина не явля-

ются основными используемыми фагами. Деградация клеточной ДНК важна для него как источник нуклеотидов, реутилизируемых при синтезе фаговой ДНК. (По составу оснований фаговая ДНК отличается от клеточной использованием гидроксиметилцитозина вместо цитозина.)

Кодируемые фагом функции синтеза ДНК в инфицированной клетке можно идентифицировать с помощью мутаций, замедляющих продукцию зрелых фагов. Насущные для фага функции идентифицируют с помощью условно-летальных мутаций, разделяемых на четыре фенотипических класса:

- Мутации, при которых синтез ДНК полностью отсутствует. Они выключают гены, продукты которых входят в состав репликативного аппарата либо участвуют в поставке предшественников (особенно гидроксиметилцитозина).
- Мутации, при которых синтез ДНК стартует с опозданием (вследствие нарушения инициации репликации). Это так называемые мутанты ДНК-задержки.
- Мутации, при которых синтез ДНК стартует, но вскоре оказывается заблокирован вследствие нарушения регуляторных функций, ДНК-лигазы и некоторых ферментов, участвующих в деградации ДНК хозяина. Это так называемые мутанты ДНК-ареста.
- Мутации, поражающие белки, участие которых в репликации не столь существенно — например, участвующие в глюкозилировании гидроксиметилцитозина ДНК.

Синтез ДНК Т4 катализируется мультиферментным комплексом, собранным из продуктов небольшой группы насущных генов фага.

Белок гена *32* (gp32) выступает в качестве SSB (см. разд. 18.7 «Наработка одноцепочечной ДНК для размножения фага в рамках модельной системы φX»), отличается высокой степенью кооперативности и необходим в больших количествах. Среди белков такого типа gp32 был охарактеризован первым. Вероятно, наличие такого белка необходимо фагу, поскольку SSB *E. coli* не подходит для репликативной вилки Т4. gp32 образует комплекс с ДНК-полимеразой фага Т4, это взаимодействие может быть важным элементом конструкции репликативной вилки.

Способ затравления синтеза ДНК системой фага Т4 тот же, что и у хозяина. С одноцепочечной ДНК Т4 в качестве матрицы белки генов *41* (p41) и *61* (p61) совместно синтезируют короткие РНК-затравки. Они ведут себя аналогично белкам DnaB и DnaG *E. coli*. Продукт гена *41* представляет собой альтернативу DnaB. Это гексамерная хеликаза, использующая энергию гидролиза ГТФ для расплетания ДНК. Комплекс p41/p61 процессивно продвигается в 5′–3′ направлении синтеза отстающей цепи, периодически инициируя фрагменты Оказаки. Другой белок, продукт гена *59*, загружает комплекс p41/p61 на ДНК; для сборки хеликазы на ДНК необходимо сместить с нее белок p32.

Продукт гена *61* необходим в количестве, гораздо меньшем, чем большинство репликативных белков Т4. На клетку приходится всего 10 копий gp61. (Это обстоятельство долгое время препятствовало его обнаружению. Он требуется системе в столь малых количествах, что не воспринимался как необходимый компонент, а считался примесью препаратов gp32.) Аналогично DnaG *E. coli*, продукт гена *61* обладает праймазной активностью. Праймаза распознает на матрице последовательность 3′-TTG-5′ и синтезирует пентарибонуклеотиды с последовательностью вида pppApCpNpNpNp. В присутствии полного репликативного аппарата эти праймеры достраиваются до цепей ДНК.

Продукт гена *43* — ДНК-полимераза с обычной 5′–3′ синтезирующей активностью, дополненной 3′–5′ экзонуклеазной корректирующей активностью. Она катализирует синтез ДНК и удаляет затравки. Скорость продвижения ДНК-полимеразы фага Т4 по одноцепочечной матрице неравномерна. Истинно одноцепочечные участки фермент проходит быстро; однако через участки с выраженной вторичной структурой матрицы фермент проходит гораздо медленнее. В прохождении этих «заторов» при сохранении скорости ДНК-полимеразе Т4 помогают «белки доступа».

Другие три белка фага Т4 так и называются: «белки доступа полимеразы». Они увеличивают сродство ДНК-полимеразы к ДНК, а также увеличивают ее скорость и процессивность. Продукт гена *45* (p45) образует тример, который работает как скользящий зажим. По своей структуре тример подобен β-димеру *E. coli* в том, что он кольцом охватывает ДНК, крепко держа ДНК-полимеразу у матрицы.

Продукты генов *44* и *62* образуют друг с другом тесный комплекс, обладающий АТФазной активностью. Это эквивалент γδ клэмп-лоудера, и его роль состоит в наложении p45 на ДНК. При наложении на ДНК p45-зажима с p43-ДНК-полимеразой расходуется (путем гидролиза) четыре молекулы АТФ.

Структура Т4 реплисомы в целом соответствует реплисоме *E. coli*. Она состоит из двух холоферментных половинок — по одной на ведущую и отстающую цепи. В данном случае димеризующиеся каталитические субъединицы (p43) взаимодействуют напрямую, а p32 принимает участие в координации действий двух ДНК-полимеразных модулей.

Пристально рассмотрев движение репликативной вилки и ее окрестности, обратим внимание на другие связанные с репликацией фага функции. Их нарушения наблюдаются у мутантов ДНК-задержки и ДНК-ареста. Три из четырех генов, мутации в которых вызывают ДНК-задержку — *39*, *52* и *60*, — кодируют три субъединицы Т4 топоизомеразы второго типа, активность которой нужна для борьбы со сверхспирализацией матрицы (см. разд. 19.13 «Топоизомеразы изменяют число сверхвитков, релаксируя или сверхспирализуя ДНК»). Возможно, какой-то этап репликации ДНК Т4 сопряжен с изменением ее топологии по сравнению с исходной линейной формой. Возможно также, что топоизомераза обеспечивает ротацию ДНК впереди репликативной вилки.

Для репликации необходим стандартный набор энзиматических активностей

Функция	*E. coli*	HeLa/SV40	Фаг T4
Хеликаза	DnaB	Т-антиген	41
Рекрутинг хеликазы/праймазы	DnaC	Т-антиген	59
Стабилизация одиночных цепей матрицы	SSB	RPA	32
Синтез затравки	DnaG	Polα/праймаза	61
Скользящий зажим	β	PCNA	45
Наложение зажима (АТФаза)	комплекс γδ	RFC	44/62
Катализ полимеризации нуклеотидов	*коровая Pol III*	Polδ	43
Димеризация холофермента	τ	?	43
Удаление РНК	*Pol I*	MF1	43
Лигирование	*ДНК-лигаза*	ДНК-лигаза 1	T4 ДНК-лигаза

РИС. 18.26 **Во всех репликативных вилках нужны одни и те же функции**

При сравнении репликативных аппаратов T4 и *E. coli* видно, как разные системы используют аналогичные пути для решения проблем, связанных с репликацией. Теперь мы можем сравнить энзиматические активности и структуры, ассоциированные с репликативными вилками *E. coli*, T4 и клеток HeLa (человека). В таблице на РИС. 18.26 перечислены функции с присвоением их конкретным белкам. Известные свойства белков репликативного комплекса можно интерпретировать в терминах сходства их функций, включая расплетание, прайминг, синтез и сшивание фрагментов. Компоненты каждой системы взаимодействуют строго очерченными способами. Это иллюстрируется тем обстоятельством, что фагу T4 требуются собственные хеликаза, праймаза, зажим и т. д., которые не могут быть заменены соответствующими бактериальными белками.

18.15 Формирование репликативных вилок в ориджине

Ключевые понятия:

- Инициация в OriC требует последовательной сборки обширного белкового комплекса.
- DnaA связывается с короткими повторами и формирует олигомерный комплекс, денатурирующий ДНК.
- Шесть мономеров DnaC связывают гексамер DnaB, и весь этот комплекс взаимодействует с местом начала репликации.
- Гексамер DnaB формирует репликативную вилку; также необходимы гираза и SSB.

Для начала цикла репликации двухцепочечной ДНК требуется несколько сменяющих друг друга активностей:

- Две цепи ДНК должны претерпеть первоначальное разделение. Это, по сути, реакция плавления на небольшом участке.
- Точка расплетания начинает двигаться вдоль ДНК; получается репликативная вилка, продолжающая движение во время элонгации.
- Первые нуклеотиды новой цепи входят в состав затравки. Ее синтез требуется один раз для ведущей цепи, а для отстающей цепи имеет место в начале каждого фрагмента Оказаки.

Таким образом, некоторые процессы, относящиеся к инициации, происходит исключительно в ориджине; другие воспроизводятся при инициации фрагментов Оказаки в ходе элонгации.

Для разработки бесклеточной экспериментальной системы репликации были использованы плазмиды, несущие последовательность *oriC E. coli*. Инициация репликации в *oriC in vitro* начинается с образования комплекса из шести белков: DnaA, DnaB, DnaC, HU, гиразы и SSB. Из шести белков, участвующих в предпрайминге, DnaA заслуживает отдельного внимания как единственный уникально задействованный в элонгации «визави». Комплекс DnaB/DnaC образует основу «машины» инициации в ориджине.

Первая стадия образования инициаторного комплекса — связывание белка DnaA с *oriC*. Эта реакция опирается на два типа повторяющихся последовательностей ДНК: в 9 пн и в 13 пн. Эти повторы совместно определяют границы минимального ориджина протяженностью 245 пн, что показано на РИС. 18.27. Ориджин активируется, когда связывание DnaA с ДНК привлекает другие белки; последовательность событий изображена на РИС. 18.28.

Консенсусные последовательности в 9 пн на правой стороне *oriC* служат сайтами первоначальной посадки мономеров DnaA, которые, кооперативно связываясь с ДНК, образуют центральную часть, на которую наматывается ДНК *oriC*. Затем DnaA воздействует на три А-Т-богатых тандемных повтора в 13 пн, расположенных на левой стороне *oriC*. В присутствии АТФ DnaA расплавляет ДНК этих сайтов с образованием открытого комплекса. Все 13 пн-повторы должны быть открыты, чтобы процесс перешел на следующую стадию.

Итак, от двух до четырех мономеров DnaA связываются с ориджином и два предпраймирующих комплекса DnaB-DnaC — по одному на каждую из вилок двунаправленной репликации. Каждый комплекс DnaB-DnaC состоит из шести мономеров DnaC, привязанных к гексамеру DnaB. Каждый комплекс DnaB-

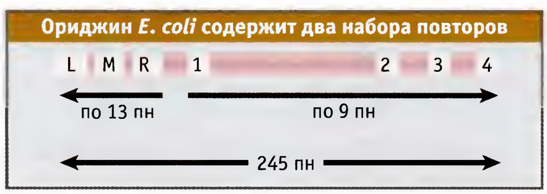


РИС. 18.27 **Минимальный ориджин соответствует интервалу между самым левым 13 пн-повтором и самым правым 9 пн-повтором**

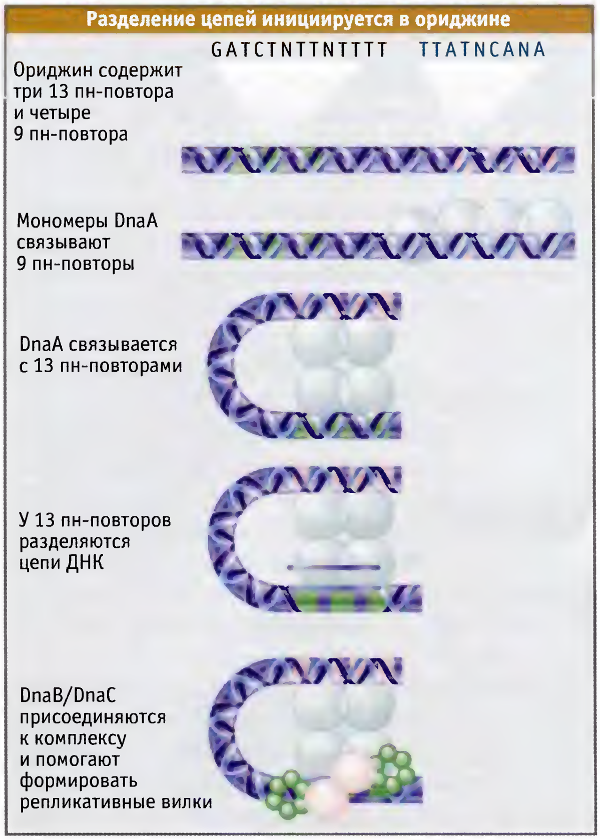


РИС. 18.28 «Предпрайминг» заключается в последовательной ассоциации белков с ориджином и образовании ими комплекса, разделяющего цепи ДНК

DnaC переносит гексамер DnaB на противоположную цепь ДНК. DnaC гидролизует АТФ для того, чтобы отпустить DnaB.

Предпраймирующий комплекс образует в *oriC* белковый агрегат массой 480 кДа, соответствующий сфере радиусом 6 нм. На электронных микрофотографиях он выглядит как темный сгусток. После начала репликации рядом со сгустком становится виден пузырек (РИС. 18.29).

Участок ДНК, на котором цепи разделены в открытый комплекс, достаточно велик, чтобы связать оба гексамера DnaB, поэтому две репликативных вилки инициируются одновременно. Когда DnaB садится, то смещает DnaA с 13 пн-повторов и увеличивает протяженность открытого участка. Затем, посредством своей хеликазной активности, DnaB увеличивает протяженность участка расплетенной ДНК. Каждый гексамер DnaB активирует одну праймазу DnaG — в одном случае, чтобы инициировать ведущую цепь; в другом — чтобы инициировать первый фрагмент Оказаки отстающей цепи.

Еще два белка требуются в поддержку реакции расплетания. Гираза служит шарниром, позволяющим одной цепи ДНК вращаться вокруг другой (эта реакция более детально обсуждается в разд. 19.15 «Гираза действует посредством инверсии витков»); без участия гиразы расплетание генерировало бы торсионные напряжения внутри молекулы ДНК. Белок SSB стабилизирует одноцепочечную ДНК по мере ее появления. Протяженность участка ДНК, расплетенного к началу репликации, вероятно, составляет <60 пн.

Белок HU *E. coli* — это ДНК-связывающий белок общего значения. Его присутствие необязательно для начала репликации *in vitro*, но стимулирует эту реакцию. HU способен сгибать ДНК и участвует в становлении открытого комплекса.

Затраты энергии в форме АТФ необходимы на нескольких этапах пред-прайминга; они также требуются для расплетания ДНК. От гидролиза АТФ зависимы и хеликазная активность DnaB, и работа гиразы. Также АТФ нужна для действия праймазы и для активации ДНК-полимеразы III.

Вслед за генерированием репликативной вилки происходит реакция прайминга, дающая начало ведущей цепи. Мы знаем, что для этого используется синтез РНК, но детали процесса неизвестны. Некоторые мутации в



РИС. 18.29 Белковый комплекс в *oriC* можно увидеть в электронный микроскоп. Визуализация проведена с использованием антител к белку DnaB. Верхняя фотография перепечатана из Funnel, B. E., et al. J. Biol. Chem. 1987. 262: 10327–10334. Copyright 1987 by American Society for Biochemistry & Molecular Biology. Фотография любезно предоставлена Barbara E. Funnell, University of Toronto. Нижняя фотография перепечатана из Barker, T. A., et al. J. Biol. Chem. 1987. 262: 6877–6885. Copyright 1987 by American Society for Biochemistry & Molecular Biology. Фотография любезно предоставлена Barbara E. Funnell, University of Toronto

dnaA могут быть супрессированы мутациями РНК-полимеразы — признак участия DnaA в этапе инициации, требующем синтеза РНК, *in vivo*.

РНК-полимераза может использоваться для синтеза затравки с примыкающих к ориджину транскрипционных единиц. Возможно, терминируя синтез РНК на сайтах ориджина, она обеспечивает свободный 3'-OH конец для затравления ДНК-полимеразы III (по аналогии с использованием D-петель в митохондриальных ориджинах — это обсуждается в разд. 15.10 «В митохондриях ориджины репликации имеют вид D-петель»).

Другой вариант объяснения роли РНК-полимеразы состоит в том, что акт транскрипции сопряжен со структурными изменениями, благоприятствующими инициации. Эту идею подкрепляет наблюдение, что транскрипция необязательно начинается в ориджине: она эффективно содействует инициации *in vitro* на расстоянии до 200 пн от ориджина и с любой цепью ДНК в качестве матрицы. Опять же *in vitro* акт транскрипции снижает необходимость сверхспирализации ДНК — еще один аргумент в пользу того, что влияние транскрипции обусловлено внесением в локальную структуру ДНК изменений, помогающих плавлению.

18.16 Сходство начала репликации у прокариот и вирусов

Ключевые понятия:

- Общий принцип инициации у бактерий состоит в том, что ориджин исходно опознается белком, который образует обширный комплекс с ДНК.
- Плавлению подвергается короткий участок А-Т-богатой ДНК.
- DnaB связывается с комплексом и формирует репликативную вилку.

Другую систему для изучения взаимодействий в ориджине представляет фаг лямбда. Карта области его ориджина показана на РИС. 18.30. Инициация репликации в ориджине лямбда требует транскрипционной активации, начинающейся в P_R. Как и для случая с *oriC*, это не обязательно подразумевает, что получающаяся при этом РНК служит затравкой для ведущей цепи. Аналогия между двумя системами подсказывает, что синтез РНК, вероятно, запускает какие-то структурные изменения в области ориджина.

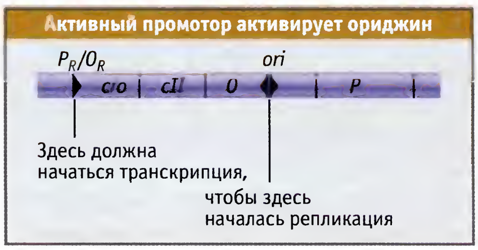


РИС. 18.30 Инициируемая в P_R транскрипция необходима для активации ориджина лямбда

Инициация требует участия продуктов генов *O* и *P* фага, а также нескольких белков хозяина. Белок O связывается с ориджином лямбда, белок P взаимодействует с белком O и с бактериальными белками. Ориджин располагается внутри гена *O*, так что белок O работает недалеко от места, где он был синтезирован.

Разновидности фага под названием λ*dv* лишены способности вызывать инфекцию. Они представляют собой короткие геномы, содержащие всю информацию, которая необходима для репликации. ДНК λ*dv* существует в бактерии как плазмида и может быть реплицирована *in vitro* при участиии белков O и P вместе с бактериальными белками репликации.

Белки O и P лямбда совместно с DnaB образуют комплекс в ориджине лямбда, *ori*λ. Ориджин состоит из двух областей, что показано на РИС. 18.31: область, заключающая в себе серию из четырех сайтов связывания белка O, соседствует с А-Т-богатой областью.

Первая ступень инициации — это связывание O с ДНК. Образуется комплекс диаметром около 11 нм, который иногда называют «O-сома». O-сома включает в себя около 100 пн (60 кДа) ДНК. В ней есть четыре сайта связывания белка O, по 18 пн каждый, что составляет около 34 кДа ДНК. Каждый сайт представляет собой палиндром, и, вероятно, связывает симметричный O-димер. По-видимому, в каждом из сайтов ДНК образует изгиб, который усиливается при связывании белка O.

Если ДНК сверхспирализована, связывание белка O вызывает в ориджине структурные изменения. Непосредственно примыкающий к O-связывающим сайтам А-Т-богатый участок становится уязвим для нуклеазы S1 — фермента, узнающего неспаренную ДНК. Это индикатор реакции плавления, которая имеет место рядом с комплексом белков O.

Роль белка O аналогична роли DnaA в *oriC*: он подготавливает ориджин к связыванию DnaB. Собственный белок лямбда, P, заменяя DnaC, привносит DnaB в ориджин. С появлением белка P фага лямбда и белка DnaB бактерии пред-праймирующий комплекс увеличивается и становится асимметричным. Он включает в себя больше ДНК (всего около 160 пн), равно как и белков. Особая роль фагового белка P — ингибирование хеликазной активности DnaB. Запуск движения репликативной вилки происходит, когда белок P покидает комплекс.

Некоторые белки важны для репликации, не участвуя непосредственно в синтезе ДНК. Интересными примерами такого рода служат DnaK и DnaJ. DnaK — шаперон, родственный обычному для эукариот стресс-белку. Его способность взаимодействовать с другими белками в конформацинно-зависимой манере играет роль во многих клеточных активностях, в том числе —

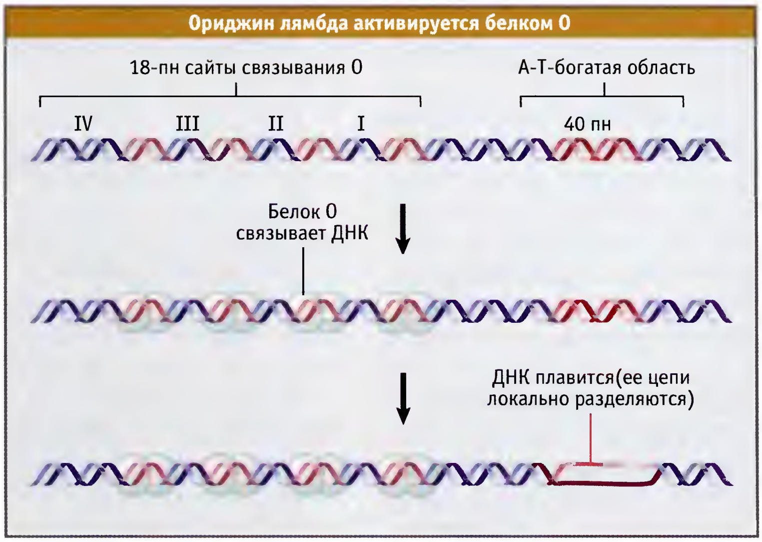


РИС. 18.31 Ориджин репликации лямбда содержит две области. Начальный этап инициации катализирует белок О, связывающий серию из четырех сайтов в левой области ориджина. Затем в смежной А-Т-богатой области происходит разделение цепей ДНК. На рисунке ДНК изображена в виде прямого дуплекса, на самом деле ориджин лямбда изогнут

в репликации. Возможно, роль DnaK и DnaJ состоит в разборке пред-праймирующего комплекса: вызывая высвобождение белка P, они «позволяют» репликации начаться.

Процессы инициации в *oriC* и в *oriλ* похожи (они включают одинаковые этапы и перекрывающиеся наборы компонентов). Первый этап состоит в узнавании ориджина белком, образующим комплекс с ДНК — DnaA для *oriC* и белок O для *oriλ*. В результате данного этапа короткий участок АТ-богатой ДНК оказывается расплавлен. После этого идет «загрузка» DnaB, что требует участия белков, отличающихся для *oriC*, *oriλ* и других ориджинов. Когда в комплекс «вступает» DnaB, образуется репликативная вилка и происходит синтез РНК-затравки, после чего начинается репликация.

События в *oriC* и в *oriλ* «вписываются» в общую модель активации ориджина. Нечто подобное происходит с ориджином вируса SV40 в клетках млекопитающих. Два гексамера Т антигена (белка, кодируемого вирусом) связываются с серией повторяющихся сайтов ДНК. В присутствии АТФ структура ДНК изменяется, что приводит к реакции плавления. В случае вируса SV40 расплавленный участок короткий и не А-Т-богатый, но отличается необычным составом азотистых оснований: одна цепь ДНК состоит почти из одних только пиримидинов, а другая — из пуринов. Неподалеку от этого сайта расположен другой важный участок. Он состоит из пар оснований А-Т, по которым ДНК согнута; посадка Т-антигена приводит к ее дальнейшему плавлению. Интересное отличие от прокариотических систем состоит в том, что Т-антиген сам обладает хеликазной активностью, обеспечивающей расплетание ДНК, и эквивалент DnaB этой системе не нужен.

18.17 Праймосома перезапускает репликацию

Ключевые понятия:

- **Инициация репликации φX требует, чтобы праймосома сместила SSB с ориджина.**
- **На поврежденном участке ДНК репликативная вилка останавливается.**
- **После устранения повреждения для возобновления репликации требуется праймосома.**
- **Белок Tus связывается с сайтами *ter* и препятствует раскручиванию ДНК белком DnaB, что приводит к завершению репликации.**

В ранних работах по репликации экстенсивное использование фага φX174 привело к открытию сложной системы затравления (праймирования). ДНК φX174 сама по себе не является субстратом для репликативного аппарата, поскольку «голая» ДНК не годится в качестве матрицы. Для начала процесса необходима одноцепочечная, покрытая белком SSB форма ДНК. Сборка **праймосомы** происходит на уникальном сайте одноцепочечной ДНК, который называется сайтом сборки праймосомы (primosome assembly site; сокращенно, *pas*). Сайт *pas* можно также назвать сайтом начала синтеза комплементарной цепи φX174. Праймосома состоит из шести белков: PriA, PriB, PriC, DnaT, DnaB и DnaC. Способность PriA смещать SSB с одноцепочечной ДНК является ключевой для посадки праймосомы на ДНК.

Несмотря на то что праймосома исходно образуется на уникальном сайте *pas* ДНК φX174, праймеры инициируются на множестве сайтов. Чтобы добраться до следующего сайта, PriA передвигается вдоль ДНК, смещая SSB. Как и в репликонах *oriC*, в репликонах φX ключевую роль

в расплетании и прайминге играет DnaB. Задача PriA — погрузка DnaB для образования репликативной вилки.

Исследователей долгое время удивлял тот факт, что в ориджине ϕX используется комплексная структура, ненужная для репликации бактериальной хромосомы. Почему бактерия обеспечивает фагу этот комплекс?

Ответ на этот вопрос кроется в застрявших репликативных вилках. На рис. 18.32 проведено сравнение движущейся репликативной вилки с тем, что происходит, если одно из оснований ДНК повреждено или в одной из ее цепей произошел разрыв (ник). И в том, и в другом случае синтез ДНК приостанавливается, а репликативная вилка застревает или разрушается. Пока неясно, остаются компоненты вилки ассоциированы с ДНК или демонтируются. Застрявшая вилка — распространенное явление; по разным оценкам, от 18 до 50% клеток *E. coli* сталкиваются в ходе репликативного цикла с этой проблемой.

Ситуацию спасает акт рекомбинации, при котором повреждение вырезается и новый дуплекс заменяет область двухцепочечного разрыва (см. рис. 20.20 в разд. 20.8 «Системы рекомбинационной репарации *E. coli*»). Принцип акта репарации основан на том, что в информационном плане комплементарные цепи ДНК дублируют друг друга. На рис. 18.33 показаны ключевые этапы акта репарации. По сути, информация с неповрежденного дочернего дуплекса используется для починки поврежденной последовательности. При этом образуется типичная для рекомбинации перекрестная связка. Работу по созданию этой связки выполняет специализированная ферментная система, специализирующаяся на гомологичной рекомбинации. По мнению ряда исследователей, главной компетенцией этой системы в

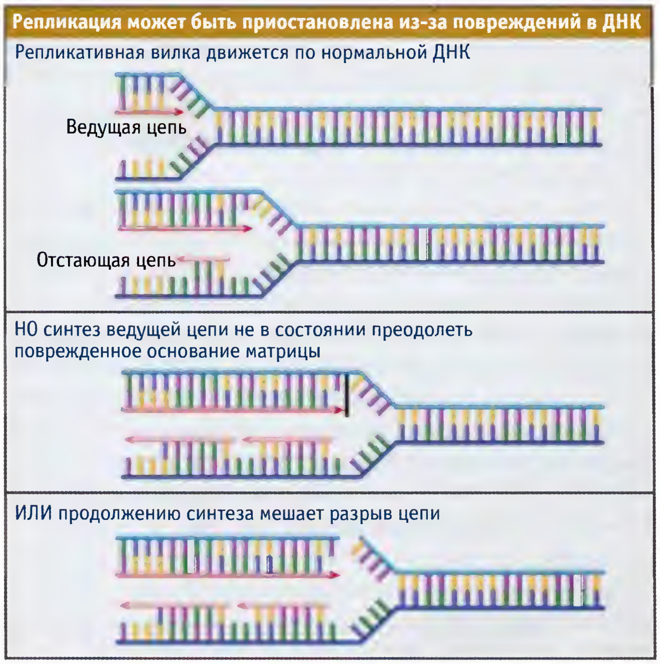


Рис. 18.32 Репликация приостановлена из-за поврежденного основания или разрыва в матричной цепи

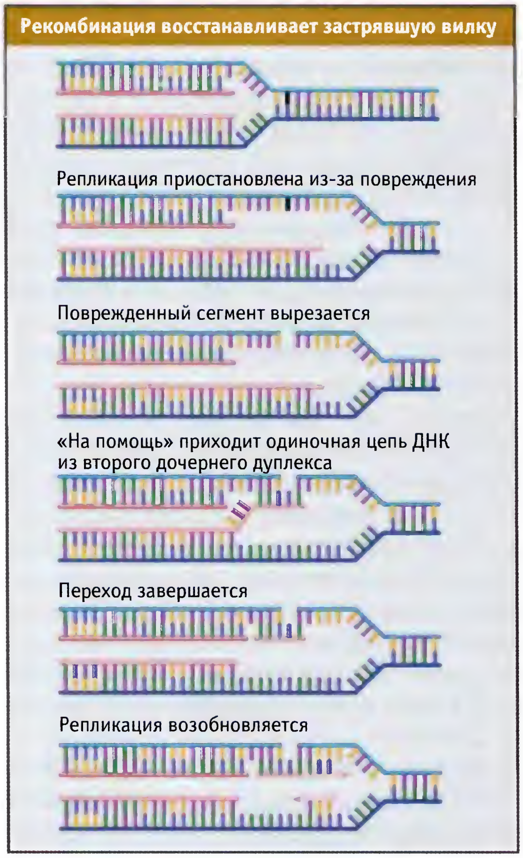


Рис. 18.33 Если дефект матрицы мешает репликации идти дальше, поврежденный сегмент ДНК может быть вырезан. Его отсутствие компенсируется актом рекомбинации с переходом цепи из второго дочернего дуплекса. После разводки гибрида дочерних дуплексов репликация возобновляется, бреши заполняются новым материалом

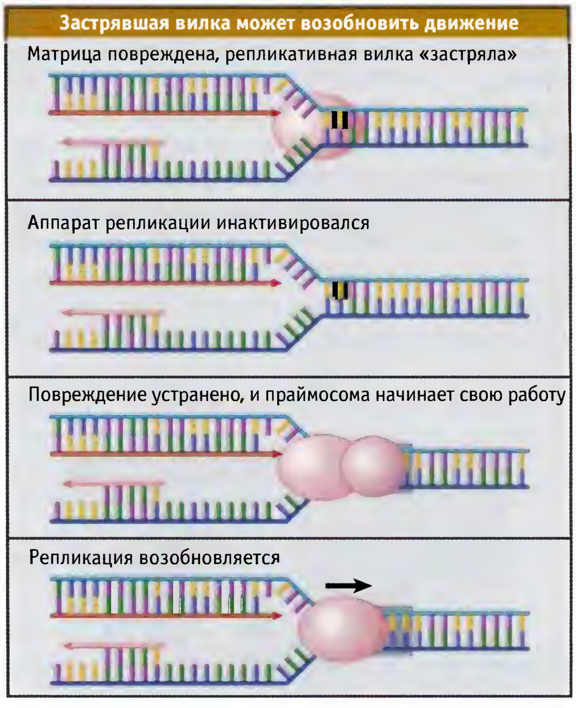


Рис. 18.34 Для перезапуска репликативной вилки после местной репарации нужна праймосома



РИС. 18.35 У *E. coli* участки терминации *ter* находятся за местом встречи репликативных вилок

клетке является именно репарация поврежденной ДНК в застрявших репликативных вилках.

После исправления повреждения репликативную вилку необходимо перезапустить. На РИС. 18.34 показано, как перезапуск выполняется посредством сборки праймосомы и «перезагрузки» DnaB, который в этом случае проявляет хеликазную активность.

Реактивация репликативной вилки — стандартная реакция (и потому важная). Потребность в ней возникает, возможно, в большинстве репликативных циклов хромосомы.

При терминации репликации вилки должны быть остановлены и демонтированы. Как это происходит?

Движение вилок останавливают специальные последовательности ДНК — *ter* элементы в хромосоме *E. coli* или их аналоги в некоторых плазмидах (см. РИС. 18.35). Общей чертой этих элементов является консенсусная последовательность в 23 пн, которая служит сайтом связывания продукта гена *tus* — 36 кДа белка, необходимого

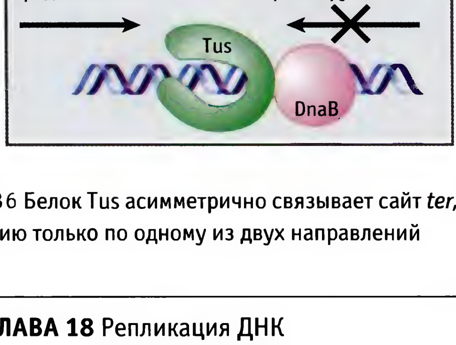


РИС. 18.36 Белок Tus асимметрично связывает сайт *ter*, блокируя репликацию только по одному из двух направлений

для терминации. Tus связывается с консенсусом и оказывает антихеликазное действие, останавливая расплетание ДНК DnaB. Ведущая цепь в продолжение своего синтеза лежит вплотную к *ter* элементу, в то время как ближайшая инициация фрагмента отстающей цепи отодвинута от *ter* на 50–100 пн.

РИС. 18.36 напоминает нам, что Tus останавливает движение репликативной вилки лишь в одном направлении. Расшифровка кристаллической структуры комплекса Tus-*ter* показывает, что белок Tus связывается с ДНК асимметрично. α-спирали белка удерживают двойную спираль с конца, что блокирует репликативную вилку. Вилка, идущая в противоположном направлении, предположительно, может сместить Tus и продолжить свое движение. Изучение системы терминации *in vivo* затруднено тем, что клетка способна справляться с ее утратой, поскольку мутации *ter* и *tus* нелетальны.

18.18 Резюме

Синтез ДНК происходит путем полунепрерывной репликации. При этом рост ведущей цепи ДНК идет непрерывно от 5′- к 3′-концу, а отстающая цепь (которая в целом растет в противоположном, 3′–5′, направлении) монтируется из коротких фрагментов Оказаки, каждый из которых синтезируется от 5′- к 3′-концу. И ведущая цепь, и каждый фрагмент Оказаки отстающей цепи инициируются РНК-затравкой (праймером), которую удлиняет ДНК-полимераза. У всех про- и эукариот есть по нескольку разновидностей этого фермента. У *E. coli* и ведущую, и отстающую цепи синтезирует ДНК-полимераза III. Работа ДНК-полимеразы III требует содействия различных белков. Некоторые из них входят в состав реплисомы, внутри которой фермент функционирует.

Реплисома, или холофермент ДНК-полимеразы III, — асимметричный димер из двух субкомплексов. Он синтезирует каждую из двух новых цепей ДНК отдельной своей половинкой. Каталитический коровый фермент (субъединица α) каждой половинки процессивно движется по матрице, удерживаемый зажимом β, который образует кольцо в охват ДНК. Зажим накладывается на ДНК комплексом клэмп-лоудера. Схожие по строению пары зажим/клэмп-лоудер широко распространены в про- и эукариотических репликативных системах.

По мере продвижения одной половины димера реплисомы вдоль матрицы ведущей цепи, другая половина протягивает через себя матрицу отстающей цепи в обратном направлении, что сопровождается образованием петли матрицы ДНК. Часть образующейся петли — синтезируемый в данный момент фрагмент Оказаки в дуплексе с матрицей. После завершения синтеза фрагмента происходит разъединение каталитической субъединицы и ДНК отстающей цепи, с

последующим прикреплением на сайте затравления следующего фрагмента Оказаки.

Белок DnaB обеспечивает хеликазную активность в репликативной вилке; его активность зависит от расщепления АТФ. В репликонах *oriC* DnaB может самостоятельно влиять на активность праймосомы, периодически взаимодействуя с DnaG — праймазой, синтезирующей РНК.

Фаг Т4 кодирует репликативный аппарат, состоящий из семи белков: ДНК-полимеразы, хеликазы, SSB ферментов затравления и «белков доступа». Подобные функции требуются и в других репликативных системах, в том числе при репликации ДНК вируса SV40 в клетках HeLa. Разные ферменты (ДНК-полимераза α и ДНК-полимераза δ, соответственно) инициируют и элонгируют новые цепи ДНК в клетках эукариот.

В системе ϕX для затравления также требуются DnaB, DnaC и DnaT. PriA — компонент, который определяет сайт сборки праймосомы (*pas*) в репликонах ϕX. PriA смещает SSB с ДНК, при этом расходуется АТФ. PriB и PriC — дополнительные компоненты праймосомы. Важность праймосомы для бактериальной клетки состоит в том, что она используется для перезапуска репликации в вилках, которые «застряли» из-за повреждений ДНК.

Общая схема активации ориджина включает в себя первоначальное локальное плавление двойной спирали, за которым следует более широкое расплетание с образованием одиночных цепей. Репликация в *oriC E. coli* инициируется, когда DnaA связывается с серией 9 пн-повторов, а затем — с серией 13 пн-повторов, где этот белок использует энергию гидролиза АТФ для разделения цепей ДНК. Предпраймирующий комплекс DnaC-DnaB смещает DnaA. DnaC высвобождается в ходе АТФ-зависимой реакции; к DnaB приходит репликаза, и репликация инициируется в форме двух вилок, стартующих в противоположных направлениях. Похожие события имеют место в ориджине фага лямбда, где в них участвуют белк О и Р фага — альтернатива бактериальных белков DnaA и DnaC, соответственно. При репликации SV40 функциональные свойства сразу нескольких участников процесса сочетает в себе Т-антиген.

Доступ DnaA в ориджин — важная составляющая системы, определяющая, когда должны начаться репликативные циклы. Помимо всего прочего, *oriC* должен конкурировать с сайтом *dat* за связывание DnaA. После того как инициация состоялась, DnaA, под действием «скользящего зажима» β, гидролизует свою АТФ, таким образом инактивируясь.

В ДНК ориджина *E. coli* есть сайты для модификации метилазой Dam, в том числе в составе 13-пн сайтов связывания DnaA. Ориджин пребывает в полуметилированном «замороженном» состоянии около 10 мин после инициации репликативного цикла. В продолжение этого периода он ассоциирован с мембраной и ре-инициация репликации репрессирована. В этой «заморозке» (секвестрировании) ориджина участвует белок SeqA, который может взаимодействовать с DnaA.

Литература

18.1 Введение

Исследование

Hirota, Y., Ryter, A., and Jacob, F. (1968). Thermosensitive mutants of *E. coli* affected in the processes of DNA synthesis and cellular division. *Cold Spring Harbor Symp. Quant. Biol.* 33, 677–693.

18.5 ДНК-полимеразы имеют общую структуру

Обзор

Hubscher, U., Maga, G., and Spadari, S. (2002). Eukaryotic DNA polymerases. *Annu. Rev. Biochem.* 71, 133–163.

Johnson, K. A. (1993). Conformational coupling in DNA polymerase fidelity. *Annu. Rev. Biochem.* 62, 685–713.

Joyce, C. M. and Steitz, T. A. (1994). Function and structure relationships in DNA polymerases. *Annu. Rev. Biochem.* 63, 777–822.

Исследование

Shamoo, Y. and Steitz, T. A. (1999). Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. *Cell* 99, 155–166.

18.7 Наработка одноцепочечной ДНК для размножения фага в рамках модельной системы ϕX

Исследование

Dillingham, M. S., Wigley, D. B., and Webb, M. R. (2000). Demonstration of unidirectional single-stranded DNA translocation by PcrA helicase: measurement of step size and translocation speed. *Biochemistry* 39, 205–212.

Singleton, M. R., Sawaya, M. R., Ellenberger, T., and Wigley, D. B. (2000). Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell* 101, 589–600.

18.9 Холофермент ДНК-полимеразы состоит из трех субкомплексов

Ссылка

Johnson, A., and O'Donnell, M. (2005). Cellular DNA replicases: components and dynamics at the replication fork. *Annu Rev Biochem.* 74, 283–315.

Исследование

Studwell-Vaughan, P. S. and O'Donnell, M. (1991). Constitution of the twin polymerase of DNA polymerase III holoenzyme. *J. Biol. Chem.* 266, 19833–19841.

Stukenberg, P. T., Studwell-Vaughan, P. S., and O'Donnell, M. (1991). Mechanism of the sliding beta-clamp of DNA polymerase III holoenzyme. *J. Biol. Chem.* 266, 11328–11334.

18.10 Зажим контролирует ассоциацию минимального фермента с ДНК

Обзор

Benkovic, S. J., Valentine, A. M., and Salinas, F. (2001). Replisome-mediated DNA replication. *Annu. Rev. Biochem.* 70, 181–208.

Davey, M. J., Jeruzalmi, D., Kuriyan, J., and O'Donnell, M. (2002). Motors and switches: AAA+ machines within the replisome. *Nat. Rev. Mol. Cell Biol.* 3, 826–835.

Исследование

Bowman, G. D., O'Donnell, M., and Kuriyan, J. (2004). Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex. *Nature* 429, 724–730.

Jeruzalmi, D., O'Donnell, M., and Kuriyan, J. (2001). Crystal structure of the processivity clamp loader gamma (gamma) complex of *E. coli* DNA polymerase III. *Cell* 106, 429–441.

Kong, X. P., Onrust, R., O'Donnell, M., and Kuriyan, J. (1992). Three-dimensional structure of the beta subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell* 69, 425–437.

18.11 Координирование синтеза ведущей и отстающей цепей

Исследование

Dervyn, E., Suski, C., Daniel, R., Bruand, C., Chapuis, J., Errington, J., Janniere, L., and Ehrlich, S. D. (2001). Two essential DNA polymerases at the bacterial replication fork. *Science* 294, 1716–1719.

18.12 Фрагменты Оказаки соединяются лигазой

Обзор

Liu, Y., Kao, H. I., and Bambara, R. A. (2004). Flap endonuclease 1: a central component of DNA metabolism. *Annu. Rev. Biochem.* 73, 589–615.

18.13 Инициация и элонгация у эукариот осуществляются разными полимеразами

Обзор

Goodman, M. F. (2002). Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu. Rev. Biochem.* 71, 17–50.

Hubscher, U., Maga, G., and Spadari, S. (2002). Eukaryotic DNA polymerases. *Annu. Rev. Biochem.* 71, 133–163.

Kaguni, L. S. (2004). DNA polymerase gamma, the mitochondrial replicase. *Annu. Rev. Biochem.* 73, 293–320.

Исследование

Karthikeyan, R., Vonarx, E. J., Straffon, A. F., Simon, M., Faye, G., and Kunz, B. A. (2000). Evidence from mutational specificity studies that yeast DNA polymerases delta and epsilon replicate different DNA strands at an intracellular replication fork. *J. Mol. Biol.* 299, 405–419.

Shiomi, Y., Usukura, J., Masamura, Y., Takeyasu, K., Nakayama, Y., Obuse, C., Yoshikawa, H., and Tsurimoto, T. (2000). ATP-dependent structural change of the eukaryotic clamp-loader protein, replication factor C. *Proc. Natl. Acad. Sci. USA* 97, 14127–14132.

Waga, S., Masuda, T., Takisawa, H., and Sugino, A. (2001). DNA polymerase epsilon is required for coordinated and efficient chromosomal DNA replication in *Xenopus* egg extracts. *Proc. Natl. Acad. Sci. USA* 98, 4978–4983.

Zuo, S., Bermudez, V., Zhang, G., Kelman, Z., and Hurwitz, J. (2000). Structure and activity associated with multiple forms of *S. pombe* DNA polymerase delta. *J. Biol. Chem.* 275, 5153–5162.

18.14 Фаг T4 обладает собственным репликативным аппаратом

Исследование

Ishmael, F. T., Alley, S. C., and Benkovic, S. J. (2002). Assembly of the bacteriophage T4 helicase: architecture and stoichiometry of the gp41-gp59 complex. *J. Biol. Chem.* 277, 20555–20562.

Salinas, F., and Benkovic, S. J. (2000). Characterization of bacteriophage T4-coordinated leading- and lagging-strand synthesis on a minicircle substrate. *Proc. Natl. Acad. Sci. USA* 97, 7196–7201.

Schrock, R. D. and Alberts, B. (1996). Processivity of the gene 41 DNA helicase at the bacteriophage T4 DNA replication fork. *J. Biol. Chem.* 271, 16678–16682.

18.15 Формирование репликативных вилок в ориджине

Исследование

Bramhill, D. and Kornberg, A. (1988). Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the *E. coli* chromosome. *Cell* 52, 743–755.

Fuller, R. S., Funnell, B. E., and Kornberg, A. (1984). The dnaA protein complex with the *E. coli* chromosomal replication origin (oriC) and other DNA sites. *Cell* 38, 889–900.

Funnell, B. E. and Baker, T. A. (1987). In vitro assembly of a prepriming complex at the origin of the *E. coli* chromosome. *J. Biol. Chem.* 262, 10327–10334.

Sekimizu, K, Bramhill, D., and Kornberg, A. (1987). ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the *E. coli* chromosome. *Cell* 50, 259–265.

Wahle, E., Lasken, R. S., and Kornberg, A. (1989). The dnaB-dnaC replication protein complex of *Escherichia coli*. II. Role of the complex in mobilizing dnaB functions. *J. Biol. Chem.* 264, 2469–2475.

18.17 Праймосома перезапускает репликацию

Обзор

Cox, M. M. (2001). Recombinational DNA repair of damaged replication forks in *E. coli*: questions. *Annu. Rev. Genet.* 35, 53–82.

Cox, M. M., Goodman, M. F., Kreuzer, K. N., Sherratt, D. J., Sandler, S. J., and Marians, K. J. (2000). The importance of repairing stalled replication forks. *Nature* 404, 37–41.

Kuzminov, A. (1995). Collapse and repair of replication forks in *E. coli*. *Mol. Microbiol.* 16, 373–384.

McGlynn, P. and Lloyd, R. G. (2002). Recombinational repair and restart of damaged replication forks. *Nat. Rev. Mol. Cell Biol.* 3, 859–870.

Исследование

Seigneur, M., Bidnenko, V., Ehrlich, S. D., and Michel, B. (1998). RuvAB acts at arrested replication forks. *Cell* 95, 419–430.

Гомологичная и сайт-специфичная рекомбинация

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

19.12 Сверхспирализация затрагивает структуру ДНК

- Сверхспирализация происходит только в закрытой ДНК, не имеющей свободных концов.
- Закрытой является ДНК, замкнутая в кольцо, или представленная линейной молекулой с концами, закрепленными белковой структурой.
- Всякая закрытая молекула ДНК характеризуется порядком зацепления, представляющим собой сумму числа витков двойной спирали и числа сверхвитков ($Lk = Tw + Wr$).
- Посредством перераспределения витков изменение структуры двойной спирали может компенсировать изменение ее скручивания в пространстве.
- Порядок зацепления может быть изменен только с помощью разрыва химических связей в ДНК и образования новых.

19.13 Топоизомеразы изменяют число сверхвитков, релаксируя или сверхспирализуя ДНК

- Топоизомеразы, изменяя порядок зацепления, разрушают в ДНК химические связи, меняют конформацию двойной спирали в пространстве и восстанавливают связи.
- Топоизомеразы первого типа, работая, рвут одну цепь ДНК; топоизомеразы второго типа действуют с внесением двухцепочечных разрывов.

19.14 Топоизомеразы разрывают цепи ДНК и снова их склеивают

- Топоизомеразы I типа вступают в ковалентную связь с одним из разорванных концов и, после вращения одной цепи вокруг другой, снова склеивают концы ДНК. Непрерывная череда химических связей «консервирует» энергию, и работа системы не требует энергозатрат.

19.15 Гираза действует посредством инверсии витков

- Гираза *E.coli* — топоизомераза второго типа, использующая энергию гидролиза АТФ для внесения в ДНК отрицательных сверхвитков.

19.16 Специализированная рекомбинация задействует специфичные сайты

- Специализированная рекомбинация опосредована взаимодействием между определенными (не обязательно гомологичными) сайтами.
- Фаг лямбда интегрирует в бактериальную хромосому посредством рекомбинации между сайтом фага и сайтом *att* на хромосоме *E.coli*.
- Фаг вырезается из хромосомы посредством рекомбинации между сайтами на концах линейного профага.
- Ген *int* фага лямбда кодирует интегразу, катализирующую реакцию интеграции.

19.17 Сайт-специфичная рекомбинация сопровождается разрывом и воссоединением

- Надрезы со смещением в 7 пн производятся в *attB* и *attP* и концы соединяются крест накрест.

19.18 Сайт-специфичная рекомбинация напоминает топоизомеразную активность

- Интегразы родственны топоизомеразам, и реакция рекомбинации напоминает действие топоизомеразы, за исключением того, что никованые цепи *разных* дуплексов склеиваются вместе.
- Реакция «консервирует» энергию, используя каталитический тирозин в ферменте для разрушения фосфодиэфирной связи и привязки к «рваному» 3'-концу.
- Две субъединицы фермента связываются с каждым из сайтов рекомбинации, и два таких димера сходятся, чтобы образовать комплекс, в котором происходит реакция переноса.

19.19 Рекомбинация фага лямбда происходит в интасоме

- Интеграция фага лямбда происходит в обширном комплексе, который, помимо прочих, содержит белок IHF хозяина.
- Реакция вырезания требует участия белков Int и Xis и узнает концы ДНК профага в качестве субстратов.

19.20 Молчащие и активные локусы переключаются при смене типа спаривания у дрожжей

- Локус *MAT*, отвечающий за тип спаривания у дрожжей, может давать генотип *MAT***a** или *MAT*α.
- Дрожжи с доминантным аллелем *HO* переключают тип спаривания.
- Аллель, находящийся в локусе *MAT*, называют активной кассетой.
- Есть еще две молчащих кассеты, *HML*α и *HMR***a**.
- Переключение происходит, если *MAT***a** заменяется *HML*α или если *MAT*α заменяется *HMR***a**.

19.21 Локус *MAT* кодирует белки-регуляторы

- В гаплоидных клетках типа α, *MAT*α активирует гены, определяющие α-специфичные функции, необходимые для спаривания, и инактивирует гены, обусловливающие тип спаривания **a**.
- В клетках типа **a** *MAT***a** не требуется.
- В диплоидных клетках продукты генов **a***1* и α*2* кооперативно репрессируют гаплоид-специфичные гены.

19.22 Молчащие кассеты в *HML* и *HMR* репрессированы

- *HML* и *HMR* репрессированы с участием элементов-сайленсеров.
- В число локусов, необходимых для поддержания репрессии *HML* и *HMR*, входят *SIR1-4*, *RAP1* и гены гистона H4.
- Связывание комплекса ORC с сайленсерами необходимо для инактивации.

19.23 Однонаправленный перенос инициируется локусом-реципиентом (*MAT*)

- Переключение типа спаривания инициируется двухцепочечным разрывом, который эндонуклеаза HO вносит в локус *MAT*.

19.24 Переключение контролируется регуляцией экспрессии HO

- Эндонуклеаза HO синтезируется в гаплоидных материнских клетках, и переключение «выливается» в новый тип спаривания обеих дочерних клеток.

19.25 Резюме

19.1 Введение

Без генетической рекомбинации не было бы и эволюции. Если бы обмен материалом между (гомологичными) хромосомами был невозможен, содержимое каждой отдельно взятой хромосомы оставалось бы зафиксированным в виде конкретных аллелей. Возникающие благоприятные мутации невозможно было бы отделить от неблагоприятных мутаций. Сегмент ДНК, служащий мишенью для неблагоприятных мутаций, по длине соответствовал бы целой хромосоме, а не отдельно взятому гену. В конце концов, накопление хромосомой делеционных мутаций воспрепятствовало бы ее функционированию.

Перетасовывая гены, рекомбинация служит инструментом разделения возникающих мутаций с последующей их проверкой в качестве отдельных элементов в разных сочетаниях аллелей. Рекомбинация работает и как средство распространения благоприятных аллелей, и как средство элиминирования неблагоприятных аллелей без удаления всех остальных генов, с которыми этот аллель ассоциирован. Таков базис для естественного отбора.

Рекомбинации, о которой сейчас шла речь, происходит между строго соответствующими последовательностями ДНК, так что ни одной пары оснований у рекомбинантных хромосом не добавляется и не теряется. В целом же выделяют три типа рекомбинации, общей чертой которых является то, что в результате дуплексы ДНК физически обмениваются материалом:

- Рекомбинацию между гомологичными последовательностями ДНК называют *общей*, или **гомологичной, рекомбинацией**. У эукариот она происходит при мейозе, как у самцов (во время сперматогенеза), так и у самок (во время оогенеза). Мы помним, что на соответствующей стадии мейоза из четырех хроматид в рекомбинации участвуют лишь две (см. разд. 2.7 «Рекомбинация — это результат физических перестроек ДНК»).
- Другой процесс лежит в основе рекомбинации между определенными парами последовательностей. Это впервые было описано у прокариот, где **сайт-специфичная (специализированная) рекомбинация** производит интеграцию фаговых геномов в бактериальную хромосому. В ней участвуют специфические последовательности фаговой и бактериальной ДНК, содержащие короткий участок гомологии. Ферменты, причастные к этому событию, действуют лишь на конкретную пару последовательностей-мишеней, вступающих в межмолекулярное взаимодействие. Межмолекулярные реакции того же рода имеют место во время деления бактерии для разделения двух мономерных кольцевых хромосом из димера, образовавшегося в результате гомологичной рекомбинации. Похожие реакции могут приводить к инверсии специфических участков бактериальной хромосомы.
- Третья разновидность рекомбинации позволяет ному сегменту ДНК быть вставленным в другой (относительно сходства их последовательнос **Перенос** — это способ перемещения определен элементов из одного места на хромосоме в дру Механизмы переноса основаны на «разрыве и соединении» цепей ДНК, что роднит их с проце ми рекомбинации (см. гл. 21 «Транспозоны», и гл «Ретровирусы и ретропозоны»).
- В ряде обстоятельств для контроля экспрессии ге используется их перестановка. При этом могут о(зовываться новые гены, подходящие для экспресс данных условиях, как в случае с иммуноглобулин Перестановка генов также может быть «ответст на» за переключение активного состояния с одн предсуществующего гена на другой, как в случае с пом спаривания у дрожжей, где последовательн в активном локусе может быть заменена другой ледовательностью из молчащего локуса. Оба т перестановок по своему механизму подобны тр позиции; на самом деле, их можно рассматривать транспозицию специального назначения.
- Еще один тип рекомбинации используется вир ми с РНК-геномом. В ходе его полимераза при с тезе РНК «перескакивает» с одной матрицы на д гую. В результате новосинтезированная моле

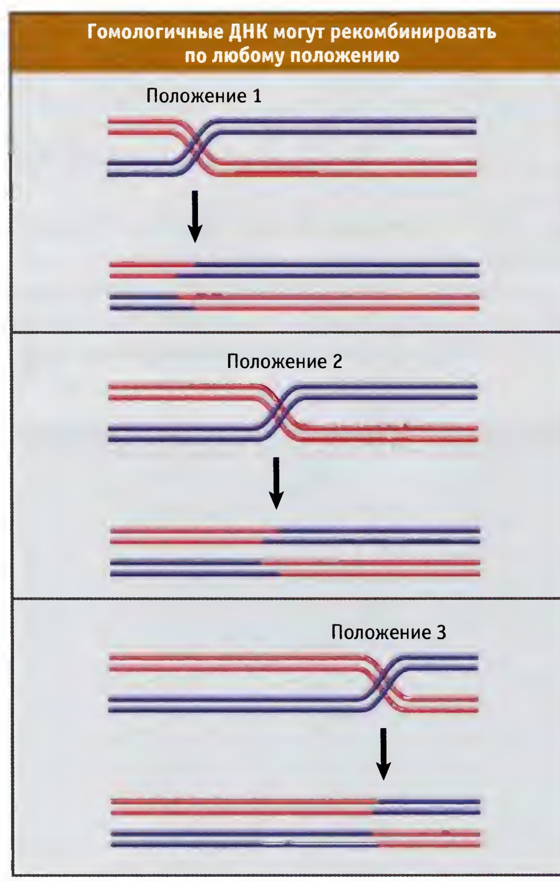


РИС. 19.1 Общая (гомологичная) рекомбинация может прои ти в любой точке двух гомологичных молекул ДНК

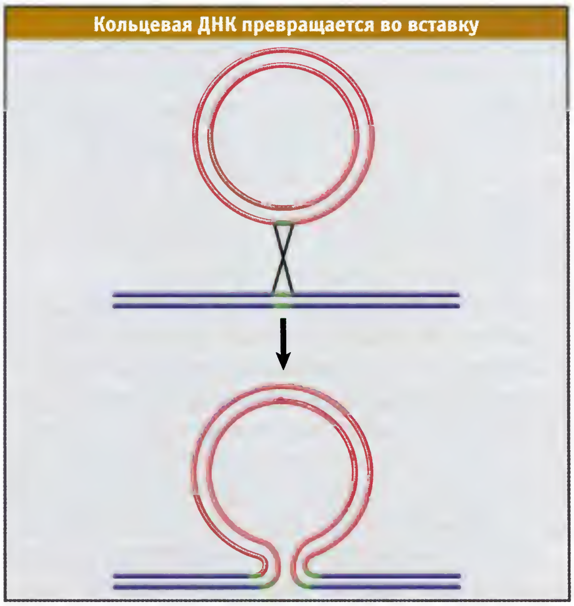


РИС. 19.2 Сайт-специфичная рекомбинация происходит между двумя конкретными гомологичными нуклеотидными последовательностями (отмеченными зеленым цветом). В остальных частях рекомбинирующие молекулы могут сколь угодно сильно различаться

сочетает в себе информацию от обоих «родителей». Такой механизм рекомбинации называют **выбором копии** и он вкратце обсуждается в разд. 22.4 «ДНК ретровируса является продуктом обратной транскрипции».

Давайте рассмотрим природу и последствия общей (гомологичной) и специализированной (сайт-специфичной) рекомбинаций.

На РИС. 19.1 уточняется, что общая рекомбинация происходит между двумя гомологичными дуплексами ДНК и может случиться в любой точке по всей их длине. Две хромосомы надрезаются в эквивалентных точках и затем снова соединяются, образуя реципрокные рекомбинанты. В точке скрещивания (оформленной на рисунке в виде X) каждая из хромосом-участниц состыкуется с второй хромосомой. При этом ДНК глобально не меняется; продукты и предшественники имеют одну и ту же структуру и гомологичны.

Специализированная рекомбинация происходит только между специфическими сайтами. Результаты зависят от местоположения двух рекомбинирующих сайтов. На РИС. 19.2 показано, как межмолекулярная рекомбинация между кольцевой и линейной ДНК вставляет кольцевую ДНК в линейную. На РИС. 19.3 показано, как внутримолекулярная рекомбинация между двумя сайтами на кольцевой ДНК высвобождает две кольцевых ДНК меньшего размера. Специализированная рекомбинация зачастую используется для такого рода организационных преобразований, обусловленных местоположением рекомбинирующих сайтов. Многое известно об осуществляющих специализированную рекомбинацию ферментах, схожих с топоизомеразами, чье действие направлено на изменение сверхспирализации ДНК.



РИС. 19.3 Сайт-специфичная рекомбинация может быть использована для превращения кольца-димера в два кольцевых мономера

19.2 Гомологичная рекомбинация происходит между хромосомами в состоянии синапсиса

Ключевые понятия:

- Хромосомы должны синапсировать (попарно сблизиться) для образования хиазм.
- Стадии мейоза могут быть соотнесены с событиями молекулярного уровня.

Гомологичная рекомбинация — это реакция между двумя дуплексами ДНК. Ферменты, причастные к ней, могут использовать в качестве субстрата любую пару гомологичных последовательностей (хотя некоторые типы последовательностей могут быть предпочтительнее других). На пространстве индивидуального генома частота рекомбинации колеблется под влиянием глобальных и локальных воздействий. Общая частота может существенно различаться для ооцитов и сперматозоидов; так, мейотическая рекомбинация происходит у женщин вдвое чаще, чем у мужчин. Внутри генома частота зависит от структуры хромосомы — в частности, обмен цепями подавляется вблизи конденсированных и неактивных участков гетерохроматина.

Гомологичная рекомбинация происходит во время длительной профазы мейоза. На РИС. 19.4 проведена параллель между видимыми изменениями хромосом (в течение пяти стадий мейотической профазы) и молекулярными взаимодействиями по обмену материалом между дуплексами ДНК.

Точка отсчета мейоза — момент, когда становятся видны отдельные хромосомы. Каждая из них предварительно прореплицировалась и состоит из двух сестринских хроматид, содержащих дуплекс ДНК. Гомоло-

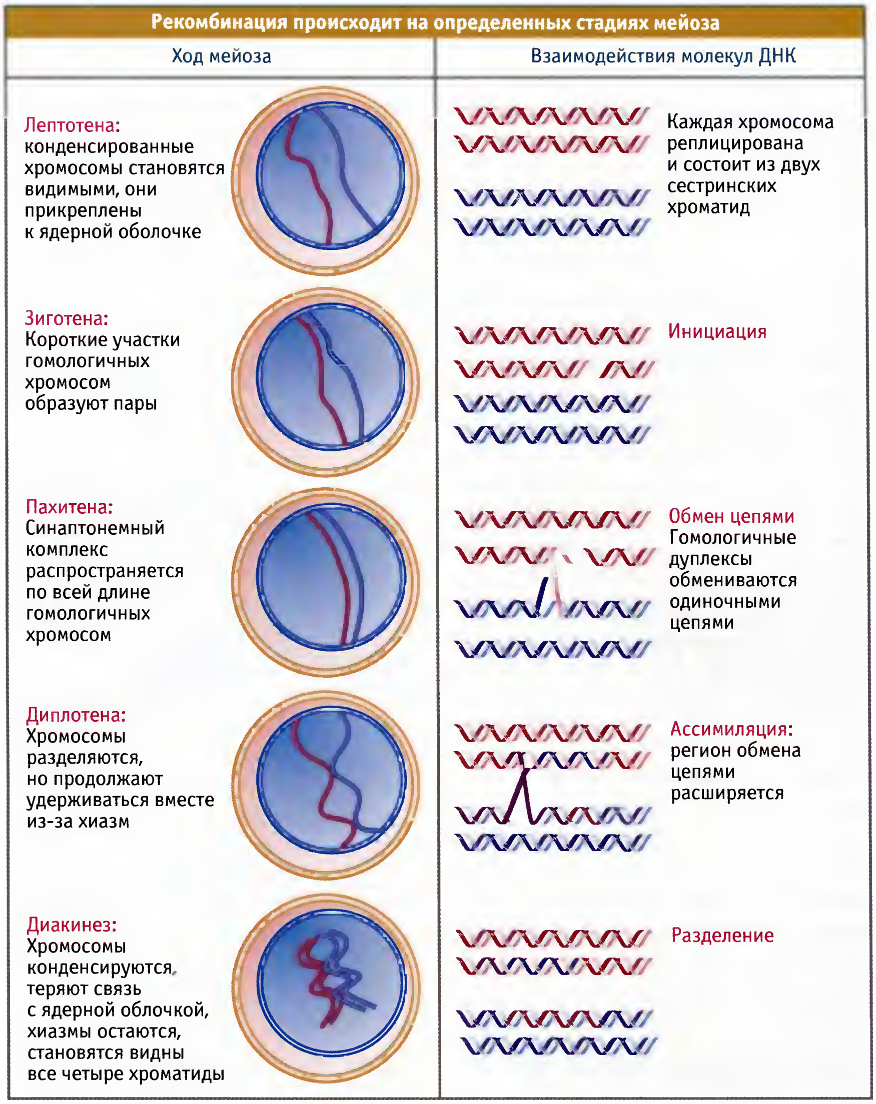


РИС. 19.4 Мейотическая рекомбинация приурочена к профазе первого деления мейоза. Стадию профазы можно определить по внешнему виду хромосом. Хотя каждая мейотическая хромосома состоит из двух реплик (сестринских хроматид), заметным это становится только в конце первой профазы. На молекулярном уровне, каждое событие кроссинговера затрагивает две из четырех гомологичных двухцепочечных молекул ДНК

гичные хромосомы подходят друг к другу и начинают спаривание по одному или более участкам, образуя **биваленты**. Спаривание продолжается до тех пор, пока каждая хромосома не противопоставится гомологу по всей своей длине. Спаривание хромосом также называют **синапсисом** (конъюгацией). По его завершении гомологичные хромосомы оказываются латерально ассоциированы в форме **синаптонемного комплекса** с характерной структурой, нюансы которой варьируют для разных видов организмов.

Рекомбинация между хромосомами — это физический обмен частями, который механистически обычно представляет собой **разрыв и воссоединение**. Две несестринских хроматиды (каждая содержит дуплекс ДНК) оказываются рассечены поперек, а затем подшиты друг к другу. Когда хромосомы начинают расходиться, можно наблюдать, как они удерживаются вместе дискретными образованиями, называемыми **хиазмами**. Число и распределение хиазм рассматривают как индикаторы генетического скрещивания на молекулярном уровне (см. рис. 19.4). В традиционном анализе принято считать соответствие между хиазмой и молекулярным скрещиванием взаимно однозначным. Хиазмы заметны и после окончательной конденсации хромосом, когда становятся видны все четыре хроматиды.

Какова же основа этих событий на молекулярном уровне? Каждая сестринская хроматида содержит один дуплекс ДНК, так что каждый бивалент содержит четыре дуплексных молекулы ДНК. Для рекомбинации требуется механизм, который позволял бы двуспиральной ДНК

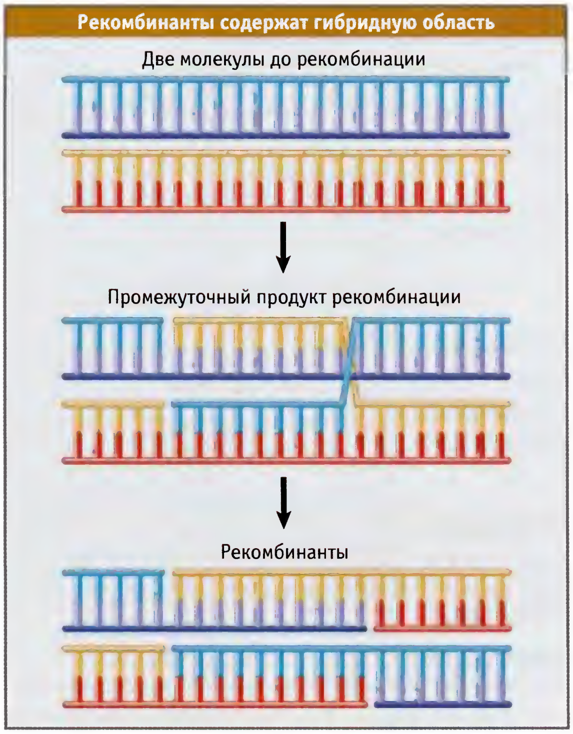


РИС. 19.5 Рекомбинация обусловлена взаимодействием комплементарных цепей из разных дуплексов

одной хроматиды взаимодействовать с двуспиральной же ДНК хроматиды другой (гомологичной) хромосомы. Эта реакция должна быть возможна для любой пары соответствующих друг другу последовательностей двух молекул в высокоспецифичной манере, позволяющей обменивать материал с высочайшей точностью, вплоть до отдельно взятой пары оснований.

Нам известен лишь один механизм опознавания нуклеиновыми кислотами последовательностей друг друга: по принципу комплементарности между одиночными цепями. На РИС. 19.5 представлена общая модель участия одиночных цепей в рекомбинации. На первом этапе в каждый дуплекс ДНК вносятся разрывы, что высвобождает одну или обе цепи дуплекса. Если хотя бы одна из цепей замещает своего гомолога в другом дуплексе, два дуплекса образуют гибрид. Этот эффект закрепляется распространением гибридной зоны, т. е. увеличением протяженности замещенного участка. Обмен обеими цепями может связать два дуплекса ДНК на манер молекулярного скрещивания, которое подлежит «разрыву и воссоединению» (то есть поступает в компетенцию соответствующих ферментных систем).

На сегодняшний день мы не в состоянии четко привязать события молекулярного уровня к изменениям на уровне хромосом. Недостаточно подробна информация о молекулярной природе рекомбинации в клетках высших эукариот (где мейоз подвергался самому пристальному наблюдению). Лишь недавно, благодаря некоторым обнаруженным у дрожжей мутациям, стало возможным соотнести ряд «молекулярных» этапов со стадиями мейоза, и то приблизительно. Существует детальная информация о процессе обмена генетическим материалом между дуплексами ДНК у бактерий, но у них рекомбинация представляет собой скорее взаимодействие между ограниченными участками, чем образование пар полноразмерными геномами. Синапсис эукариотических хромосом продолжает оставаться самой труднообъяснимой на молекулярном уровне частью процесса.

19.3 Разрыв и воссоединение с участием гетеродуплексной ДНК

Ключевые понятия:

- Ключевым событием рекомбинации между двумя дуплексами ДНК является обмен одиночными цепями.
- Когда цепь ДНК одного дуплекса вытесняет своего «двойника» из другого дуплекса, образуется разветвленная структура.
- Такой обмен генерирует на хромосоме гетеродуплексный участок, содержащий по одной цепи ДНК от каждого из исходных дуплексов.
- Два реципрокных обмена необходимы для образования соединенной молекулы.
- Соединенная молекула разводится на два обособленных дуплекса посредством разрыва двух связывающих цепей.
- Получатся рекомбинанты или нет, зависит от того, в какие цепи ДНК при разделении будут внесены разрывы: в те, что участвовали в реципрокном обмене, или в другую пару.

Акт соединения двух дуплексных ДНК лежит в основе всего процесса рекомбинации. Посему наш молекулярно-биологический анализ этого явления начнется с расширения представлений об использовании спаривания азотистых оснований между комплементарными одиночными цепями, участвующими в рекомбинации. Целесообразно представить себе реакцию рекомбинации в терминах обмена одиночными цепями (хоть мы и увидим, что она необязательно инициируется только так), поскольку свойства созданных таким образом молекул существенны для понимания процессов, составляющих основу рекомбинации.

РИС. 19.6 иллюстрирует процесс, начинающийся с разрывов в идентичных точках гомологичных цепей двух спаренных ДНК-дуплексов. Разрыв позволяет образовавшимся свободным концам двигаться. Каждая цепь оставляет своего партнера и совершает спаривание (кроссинговер) со своим комплементом в другом дуплексе.

Реципрокный обмен цепями закрепляет связь между дуплексами ДНК. Связанную таким образом пару дуплексов называют **соединенной молекулой**. Точку, в которой отдельная цепь ДНК переходит с одного дуплекса на другой, называют **рекомбинантным соединением**.

Там, где идет рекомбинация, в составе каждого участвующего в ней дуплекса ДНК есть участок, содержащий по одной цепи от каждого участвующего в ней дуплекса

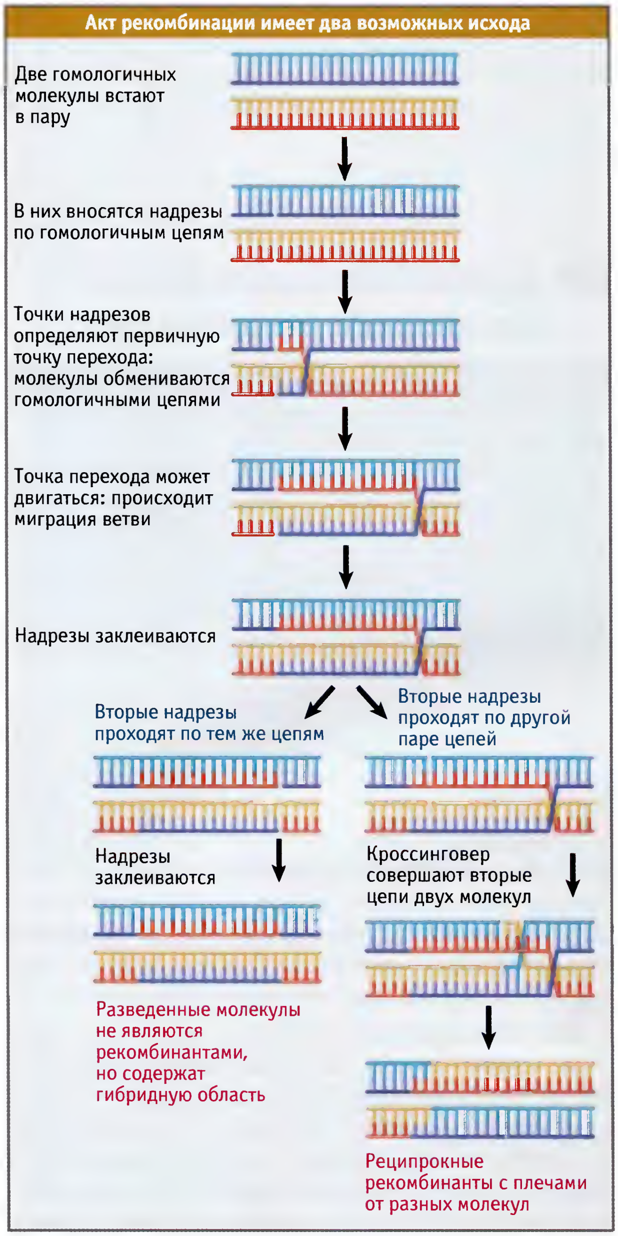


РИС. 19.6 Акт рекомбинации двух гомологичных двухцепочечных молекул включает в себя этапы внесения надрезов, кроссинговера одиночных цепей и миграции ветви

ДНК. Этот участок называется **гибридной ДНК**, или **гетеродуплексной ДНК**.

Рекомбинантное соединение отличается тем, что может двигаться вдоль дуплекса. Такую мобильность называют **миграцией ветви**. РИС. 19.7 иллюстрирует миграцию одиночной цепи в дуплексе. Точка ветвления может мигрировать в любом направлении по мере замещения одной цепи другой.

Миграция ветви важна для нас по причинам как теоретического, так и практического свойства. В фундаментальном плане, она придает рекомбинирующим структурам динамичность. С практической точки зрения, из-за нее установить *in vitro* исходную точку ветвления невозможно, поскольку ветвь могла мигрировать уже после изоляции молекулы.

Миграция ветви в промежуточном продукте рекомбинации позволяет рекомбинантному соединению двигаться в любом из двух направлений. Скорость самопроизвольной миграции ветви точно не определена, но по оценкам *in vitro*, вероятно, недостаточна, чтобы обеспечить образование протяженных участков гетеродуплексной ДНК в естественных условиях. Следовательно, любая экстенсивная миграция ветви *in vivo* должна катализироваться ферментом рекомбинации.

Соединенная молекула — продукт обмена цепями — должна быть *разделена* на две обособленных дуплексных молекулы. **Разделение** требует еще пары одноцепочечных разрывов. Самый легкий способ разобраться в возможных исходах — рассмотреть соединенную молекулу в одной плоскости как **структуру Холлидея**. В этом поможет РИС. 19.8, где изображена та же структура, что и на рис. 19.6, но с одним из дуплексов, повернутым относительно другого. Исход реакции зависит от того, какая пара цепей будет перерезана.

Если разделена будет пара цепей, которая не инициировала обмен цепями, получится, что все четыре исходных цепи были в какой-то момент перерезаны. Так высвобождаются **сплайс-рекомбинантные** молекулы ДНК. Дуплекс ДНК одного родителя ковалентно привязан к дуплексу ДНК другого родителя через участок гетеродуплексной ДНК. Между маркерами, расположенными по обе стороны гетеродуплексного участка, зафиксирован стандартный акт рекомбинации.

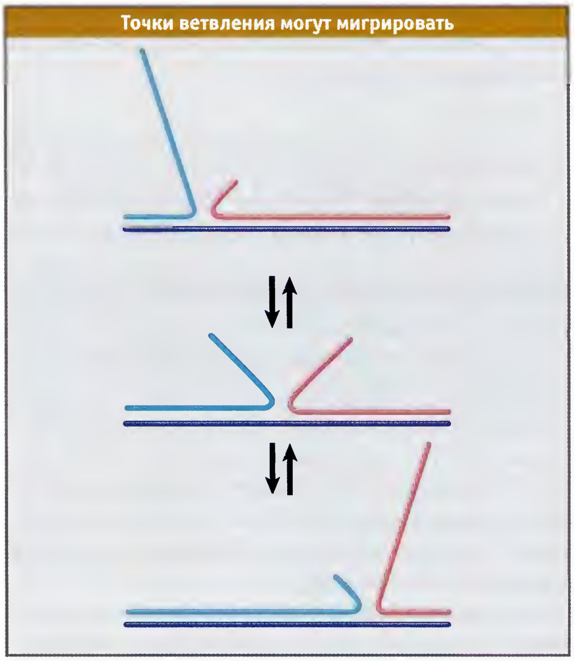


РИС. 19.7 Миграция ветви может идти в любом из двух направлений: две идентичных одиночных цепи разделяют комплементарную цепь

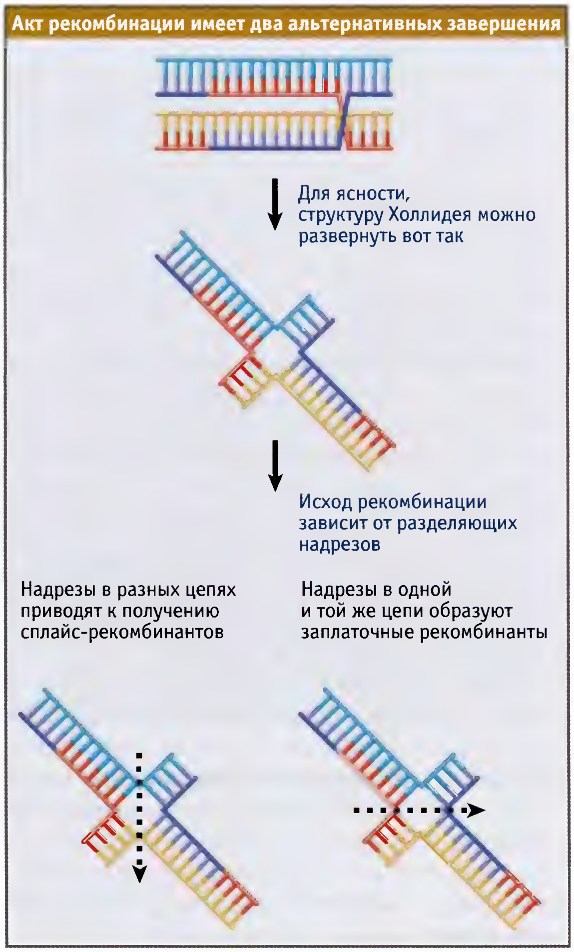


РИС. 19.8 Разделение структуры Холлидея может породить рекомбинантов или вернуть исходные молекулы, в зависимости от того, по каким цепям пройдут разделяющие надрезы. Однако вследствие миграции ветви, любой продукт разделения будет содержать гибридный участок

Если провести надрезы структуры Холлидея по тем двум цепям, которые уже были разрезаны, другие две остаются нетронутыми. Такое разделение высвобождает исходные (родительские) дуплексы почти нетронутыми, за исключением небольшого участка гетеродуплекса, остающегося после события рекомбинации. Такие молекулы называют **заплаточными рекомбинантами** (patch recombinants).

Эти альтернативные разрешения соединенной молекулы знаменуют следующий принцип: *обмен цепями между дуплексами ДНК всегда оставляет след в виде участка гетеродуплексной ДНК, однако этот обмен может и не сопровождаться рекомбинацией флангов гетеродуплексного участка.*

Какова минимальная длина участка одиночной цепи, требуемая для установления связи между рекомбинирующими дуплексами? В экспериментах по интродукции в бактерию коротких гомологичных последовательностей (с плазмидой или фагом в качестве носителя) частота рекомбинации резко падала для участков гомологии короче 75 пн — расстояния, ощутимо большего, чем десяток пн, необходимый для ассоциации комплементарных одноцепочечных участков. Это означает, что для рекомбинации требуется нечто большее, чем ренатурация комплементов как таковая.

19.4 Двухцепочечные разрывы инициируют рекомбинацию

Ключевые понятия:

- Рекомбинация инициируется двухцепочечным разрывом, вносимым в дуплекс-реципиент.
- Действие экзонуклеазы генерирует 3'-одноцепочечные концы, которые вторгаются в другой дуплекс-донор.
- Новый синтез ДНК восстанавливает разрушенный материал.
- Так образуется объединенная рекомбинантная молекула, в которой два дуплекса ДНК соединены участком гетеродуплекса.

Общая модель рекомбинации на рис. 19.4 показывает, как в одном из дуплексов должен быть произведен разрыв, чтобы одиночные цепи могли стать доступными для участия в генетическом обмене. На рис. 19.6 приведена модель, согласно которой разрывы в одиночных цепях происходят независимо и последовательно. На самом деле, обмен генетическим материалом инициируется **двухцепочечным разрывом**. Соответствующую этому утверждению модель иллюстрирует РИС. 19.9.

Рекомбинацию инициирует эндонуклеаза, разрезающая один из дуплексов-партнеров (реципиент). Действие экзонуклеазы превращает разрез в пробел. Экзонуклеаза (их может быть несколько) «обкусывает» цепь ДНК по каждую сторону разреза, оставляя одноцепочечные «хвосты», оканчивающиеся на 3'-OH. Один из этих свободных 3'-концов затем вторгается в гомологичную область другого дуплекса (донора). Этот процесс называют вторжением, или инвазией, одиночной цепи. Образование гетеродуплекса ДНК сопряжено с появлением так называемой D-петли, в которую «уходит» вытесняемая цепь дуплекса-донора. Эта D-петля растет благодаря репаративному синтезу ДНК, при котором вторгшийся свободный 3'-конец используется в качестве затравки.

Через некоторое время D-петля становится так велика, что достигает размера пробела в хроматиде-реципиенте. Когда вытесняемая из донора одиночная цепь D-петли достигает дальнего (левого) края пробела со свободным («липким») 3'-концом, их комплементарные последовательности взаимодействуют. Теперь пробел, фланкированный участками гетеродуплексной ДНК, заполнен одноцепочечной D-петлей.

Целостность дуплекса на поврежденном участке реципиента может быть восстановлена посредством репаративного синтеза с использованием 3'-конца на левой

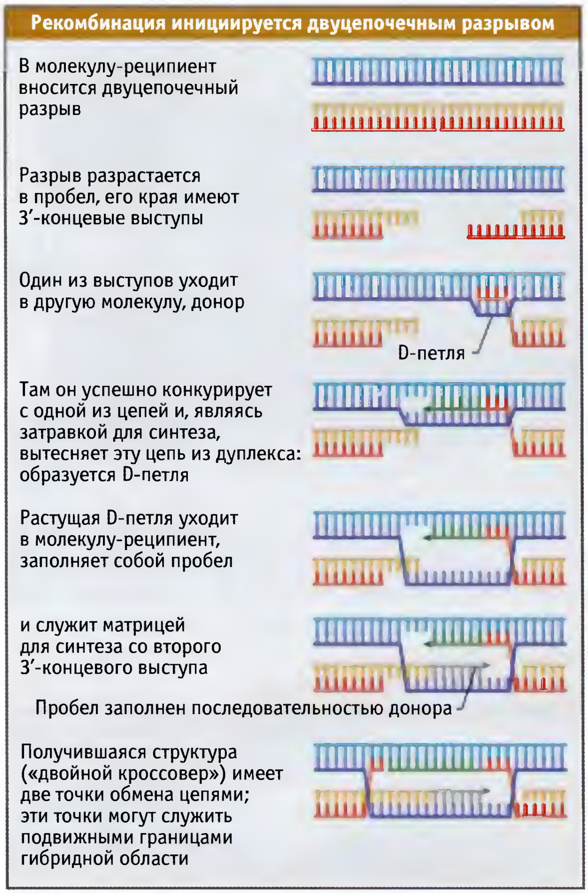


РИС. 19.9 Рекомбинация инициируется двухцепочечным разрывом; он разрастается в пробел с липкими краями, один из которых начинает вторгаться в гомологичный дуплекс

стороне пробела в качестве затравки и D-петлей в качестве матрицы. В итоге репарация пробела занимает два отдельных раунда синтеза одноцепочечной ДНК.

Миграция ветвей преобразует получившуюся структуру в двухдуплексную молекулу с двумя рекомбинантными соединениями. Для разделения соединяющие цепи должны быть перерезаны.

Если соединения надрезаются по одним и тем же цепям, то высвобождаются исходные некроссоверные молекулы, каждая — с участком измененной генетической информации, как след перенесенной рекомбинации. Если перерезаются противоположные пары цепей, результат представляет собой генетический кроссовер.

Структура двухсвязной молекулы до разводки служит иллюстрацией принципиальной разницы между моделью с двухцепочечным разрывом и моделями, опирающимися лишь на обмен одиночными цепями (модели одной цепи).

- Как следствие двухцепочечного разрыва, по оба края участка, вовлеченного в обмен цепями, образуются гетеродуплексы ДНК. Между этими сегментами гетеродуплекса находится участок, который у реципиента был разрушен (см. рис. 19.9). Теперь он (у обеих рекомбинировавших молекул) по последовательности соответствует ДНК донора. Таким образом, расположение гетеродуплексных последовательностей асимметрично, а часть одной молекулы была конвертирована в последовательность другой. Вот почему одну из хроматид — инициирующую — называют реципиентом.
- В вариантах «моделей одной цепи», где участок ДНК деградирует и синтезируется заново, инициирующая хроматида, наоборот, служит донором генетической информации. Как следствие реципрокного обмена одиночными цепями, в каждом дуплексе ДНК гетеродуплексный материал покрывает участок от сайта обмена до мигрирующей ветви (рис. 19.6).

Модель с двухцепочечным разрывом не преуменьшает важности образования гетеродуплексной ДНК, которое пребывает единственно возможным средством взаимодействия между двумя дуплексными молекулами. «Перекладывание ответственности» инициировать рекомбинацию с одно- на двухцепочечные разрывы, между тем, несколько меняет наш взгляд на то, как клетка «управляется» с собственной ДНК.

Участие двухцепочечных разрывов сначала кажется неожиданным. Как только молекула ДНК оказалась рассечена надвое, пути назад нет. Сравните события, изображенные на рис. 19.6 и 19.9. В «модели одной цепи» нигде не происходит никакой потери информации. В модели с двухцепочечным разрывом потеря информации незамедлительно следует за инициирующим разрезом. Любая ошибка в восстановлении этой информации может оказаться фатальной. С другой стороны, сама способность к восстановлению утраченной информации путем копирования ее с другого дуплекса — это важнейший для клетки страхующий фактор.

19.5 Рекомбинирующие хромосомы соединены синаптонемным комплексом

Ключевые понятия:

- В начале мейоза гомологичные хромосомы образуют пары в синаптонемном комплексе.
- Хроматиновые массы гомологов разделяет белковый комплекс.

В основе рекомбинации заложен тот парадокс, что родительские хромосомы, на первый взгляд, никогда не оказываются в контакте, достаточно близком для того, чтобы рекомбинация ДНК все-таки произошла. Хромосомы вступают в мейоз в форме реплицированных пар (сестринских хроматид), которые воспринимаются глазом как хроматиновая масса. Они спариваются, образуя **синаптонемный комплекс**, и многие годы считалось, что эта стадия как-то связана с рекомбинацией — возможно, как необходимое вступление к

обмену материалом ДНК. Более свежая точка зрения состоит в том, что синаптонемный комплекс — это, скорее, последствие, нежели причина рекомбинации. Остается определить, как структура синаптонемного комплекса соотносится с контактами между молекулами ДНК.

Синапсис начинается, когда каждая хромосома (пара сестринских хроматид) конденсируется вокруг структуры (очевидно, белковой), называемой **аксиальным элементом**. Аксиальные элементы соответствующих хромосом затем сопоставляются, образуя синаптонемный комплекс в форме трехчастной структуры, в составе которой аксиальные элементы, называемые теперь **латеральными элементами**, отделены друг от друга **центральным элементом**. Пример представлен на РИС. 19.10.

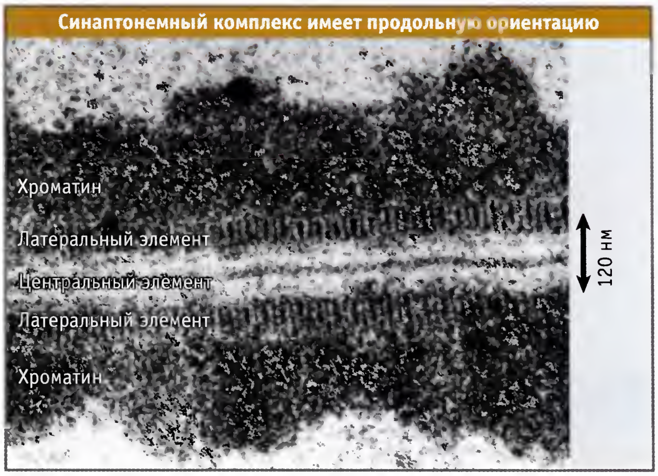


РИС. 19.10 Синаптонемный комплекс приводит хромосомы в стыковочную позицию. Воспроизведено по D. von Wettstein. 1971. *Proc. Natl. Acad. Sci. USA*. 68: 851–855. Фотография любезно предоставлена D. von Wettstein, Washington State University

На изображенной стадии каждая хромосома представлена массой хроматина, ограниченной латеральным элементом. Два латеральных элемента разделяет тонкий, но плотный, центральный элемент. Триплет параллельных плотных тяжей лежит в одной плоскости, которая изгибается и закручивается вдоль своей оси. Расстояние между гомологичными хромосомами, в молекулярных масштабах, порядочное: более 200 нм, при диаметре ДНК в 2 нм. Таким образом, главная проблема с пониманием роли синаптонемного комплекса состоит в том, что, работая «совместителем» гомологичных хромосом, он при этом отнюдь не приводит гомологичные молекулы ДНК в контакт друг с другом.

Видимую связь двух сторон синаптонемного комплекса удалось наблюдать лишь у грибов и насекомых. Сферические или цилиндрические связки, расположенные поперек комплекса, получили название **узлов**, или **рекомбинационных узелков**; они встречаются с той же частотой и распределены так же, как хиазмы. В их названии отражена надежда, что они-то и окажутся сайтами рекомбинации.

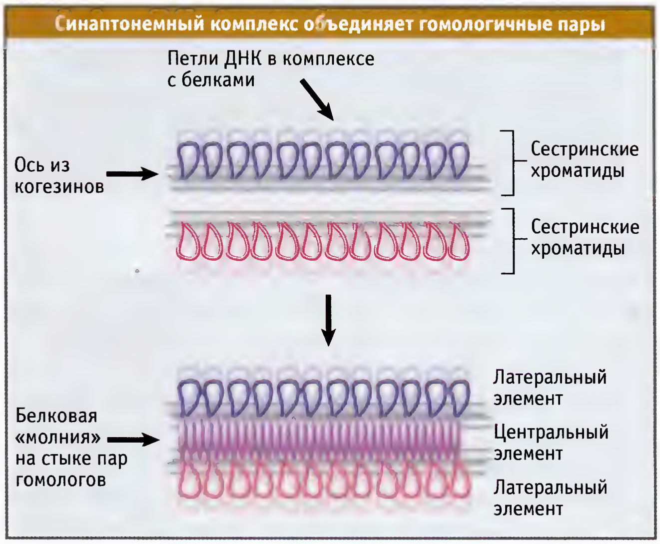


РИС. 19.11 Каждая пара сестринских хроматид имеет ось из когезинов с отходящими от нее петлями. Две таких оси, склеенных белками Zip, образуют синаптонемный комплекс

Исходя из мутаций, нарушающих образование синаптонемного комплекса, можно установить причастность конкретных типов белков к его структуре. На РИС. 19.11 представлено молекулярное видение синаптонемного комплекса. Характерные черты его структуры определяются двумя типами белков:

Когезины образуют для каждой пары сестринских хроматид одиночную прямолинейную ось, от которой «растут» хроматиновые петли. Эта структура — эквивалент латерального элемента с рис. 19.10. Когезины принадлежат к общей группе белков соединения сестринских хроматид, причастных к контролю их сегрегации при митозе или мейозе.

Латеральные элементы соединены поперечными филаментами, которые эквивалентны центральному элементу с рис. 19.10. Они образованы белками Zip.

Мутации, поражающие белки, необходимые для формирования латеральных элементов, обнаруживаются в генах, кодирующих когезины. Среди когезинов, участвующих в мейозе, — Smc3p (который используется также при митозе) и Rec8p (который специфичен для мейоза и родственен митотическому когезину Scc1p). Когезины, по-видимому, связываются со специфическими сайтами вдоль хромосомы — как при митозе, так и при мейозе. Скорее всего, они играют структурную роль в сегрегации хромосом. В мейозе образование латеральных элементов может быть необходимо для поздних стадий рекомбинации, потому что, хотя мутации в когезинах не препятствуют образованию двухцепочечных разрывов, они-таки блокируют образование рекомбинантов.

У мутантов по *zip1* латеральные элементы образуются и продольно совмещаются, но в близкий синапсис не вступают. У белка Zip1 N-концевой домен находится в

центральном элементе, а С-концевой — в латеральных элементах. Два других белка, Zip2 и Zip3, локализуются так же, как Zip1. Группа белков Zip образует поперечные филаменты, которые связывают латеральные элементы пар сестринских хроматид.

19.6 Синаптонемный комплекс образуется после внесения двухцепочечных разрывов

Ключевые понятия:

- **Двухцепочечные разрывы, которые инициируют рекомбинацию, предшествуют образованию синаптонемного комплекса.**
- **Если рекомбинация заблокирована, синаптонемный комплекс не образуется.**

Как достоверно показано для дрожжей, двухцепочечные разрывы инициируют и гомологичную, и сайт-специфичную рекомбинацию. Так, двухцепочечные разрывы вовлечены в изменение типа спаривания, при котором имеет место замена одной последовательности на другую (см. разд. 19.23 «Однонаправленный перенос инициируется локусом-реципиентом (*MAT*)»). Также двухцепочечные разрывы возникают на ранних этапах мейоза, в сайтах горячих точек рекомбинации. Их местоположение не соответствует каким-то особым типам последовательностей. Чаще всего «горячие точки» находятся в районах промоторов и вообще в областях хроматина, отличающихся хорошей доступностью. Частота рекомбинации убывает по градиенту в одну или в обе стороны от горячей точки. Горячая точка тожественна сайту, по которому рекомбинация инициируется, а градиент отражает разброс рекомбинационных процессов по ДНК и их вероятностную сущность.

Теперь мы можем истолковать роль двухцепочечных разрывов в молекулярных терминах. Концы, образующиеся при двухцепочечном разрыве, быстро преобразуются в длинные одноцепочечные 3′-концы, как показано в модели на рис. 19.9. Дрожжи, мутантные по *rad50*, у которых блокировано преобразование «обрубка» в одноцепочечный «отросток», дефектны по рекомбинации. Это признак того, что двухцепочечные разрывы необходимы для рекомбинации. Вышеупомянутый градиент определяется убывающей по мере удаления от точки разрыва вероятностью того, что экзонуклеаза укоротит ДНК до этого места, и на этом месте будет одиночная цепь.

У мутантов по *rad50* 5′-концы двухцепочечных разрывов соединены с белком Spo11, который гомологичен каталитическим субъединицам топоизомераз II. Возможно, Spo11 — это топоизомеразоподобный фермент, генерирующий двухцепочечные разрывы. Модель такой реакции представлена на РИС. 19.12. В ней допускается, что Spo11 обратимо взаимодействует

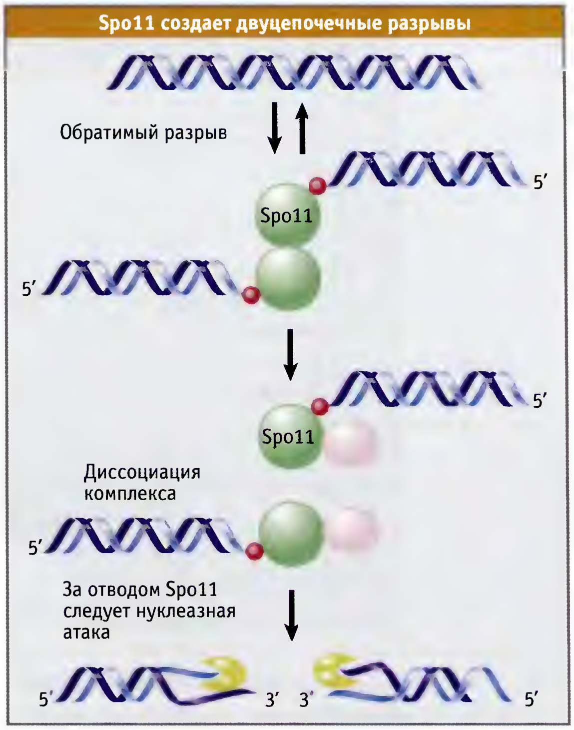


РИС. 19.12 Spo11 ковалентно пришит к 5′-концам в двухцепочечных разрывах

с ДНК; разрыв преобразуется в долговременную структуру взаимодействием с другим белком, который диссоциирует комплекс Spo11-ДНК. За устранением Spo11 следует действие нуклеазы. Еще как минимум девять разных белков требуется для процессирования двухцепочечных разрывов. Одна группа белков требуется для преобразования двухцепочечного разрыва в пару 3′-ОН одноцепочечных концов. Другая группа белков придает одноцепочечным концам способность к вторжению в гомологичный дуплекс ДНК.

Корреляция между рекомбинацией и образованием синаптонемного комплекса достоверно установлена и, как показали работы последних лет, все мутации (у *Drosophila* или у дрожжей), отменяющие образование хромосомами пар, препятствуют также и рекомбинации. Система порождения двухцепочечных разрывов, инициирующих рекомбинацию, отличается всеобщей консервативностью. Гомологи Spo11 были идентифицированы у ряда высших эукариот, а мутация по соответствующему гену у *Drosophila* блокирует всю мейотическую рекомбинацию.

Существует несколько живых систем, для которых при рекомбинации можно провести сравнение событий молекулярного и цитологического уровней. Недавний прогресс в анализе мейоза у *Saccharomyces cerevisiae* резюмирован на РИС. 19.13, где представлена временная развертка событий.

Двухцепочечные разрывы появляются, а затем исчезают, в течение 60 мин. Первые соединенные молекулы (как полагают — промежуточные продукты рекомбина-



РИС. 19.13 Двухцепочечные разрывы в ДНК появляются одновременно с аксиальными элементами хромосом; двухцепочечные разрывы исчезают во время становления синаптонемных комплексов. Соединенные молекулы ДНК детектируются, начиная с зиготены, вплоть до появления первых рекомбинантов в конце пахитены

ции) появляются вскоре после исчезновения двухцепочечных разрывов. Судя по последовательности событий, двухцепочечные разрывы, индивидуальные реакции образования хромосомных пар и формирование рекомбинантных структур сменяют друг друга в одном и том же месте на хромосоме.

Двухцепочечные разрывы появляются в продолжение периода формирования аксиальных элементов. Они исчезают во время превращения хромосом в синаптонемные комплексы. Такая раскладка событий по времени друг относительно друга соответствует представлению о том, что синаптонемный комплекс — результат инициации рекомбинации посредством внесения двухцепочечных разрывов и перехода от них к более поздним промежуточным продуктам рекомбинации. Это представление подкреплено наблюдением, что мутанты по *rad50* не могут превращать аксиальные элементы в синаптонемные комплексы. Это опровергает традиционное видение синаптонемного комплекса как символа необходимости образования хромосомами пар в преддверии молекулярной рекомбинации.

Сложно определить, происходит ли рекомбинация на стадии синапсиса, поскольку о рекомбинации судят по возникновению рекомбинантов после завершения мейоза. Отслеживая возникновение рекомбинантов у дрожжей прямо по продукции молекул ДНК, содержащих диагностические сайты рестрикции, удалось показать, что рекомбинанты появляются в конце пахитены. Это отодвигает во времени завершение рекомбинации после формирования синаптонемных комплексов.

Таким образом, синаптонемный комплекс образуется после двухцепочечных разрывов, которые инициируют рекомбинацию, и продолжает существовать до формирования рекомбинантных молекул. Маловероятно, что он необходим для рекомбинации как таковой, потому как некоторые мутанты, лишенные нормального синаптонемного комплекса, могут порождать рекомбинантов. А вот мутанты, у которых рекомбинация полностью подавлена, синаптонемный комплекс не образуют. Похоже, синаптонемный комплекс образуется как следствие начавшейся рекомбинации, вслед за разбиением хромосом на пары, и требуется для последующих стадий мейоза.

С тех пор, как была предложена модель рекомбинации через образование структуры Холлидея, предполагалось, что разделение этой структуры дает начало либо некроссоверным продуктам, либо кроссоверным рекомбинантам — в зависимости от того, какие цепи участвовали в разделении (см. рис. 19.8). Недавние измерения времен продукции кроссоверов и некроссоверов заставляют в этом усомниться. Кроссоверы впервые появляются сильно после первого появления соединенных молекул, а некроссоверы — почти одновременно с соединенными молекулами (см. рис. 19.13). Если бы оба типа рекомбинантов являлись продуктами одного и того же процесса разделения, им следовало бы появляться в одно и то же время. Наблюдаемое расхождение по времени может означать, что кроссоверы образуются, как прежде и думали, разделением соединенных молекул, в то время как для продукции некроссоверов может существовать какой-то другой путь.

19.7 Спаривание хромосом и образование синаптонемного комплекса происходят независимо друг от друга

Ключевые понятия:

- Мутации, нарушающие ход спаривания хромосом или образование синаптонемного комплекса, не влияют на другой.

Дискриминировать процессы спаривания хромосом и образования синаптонемного комплекса помогают эффекты двух мутаций, каждая из которых блокирует один из этих процессов, не нарушая другого.

При мутации гена в *zip2* хромосомы становятся в пары, но синаптонемных комплексов не образуют. Значит, узнавание гомологами друг друга происходит независимо от рекомбинации или образования синаптонемного комплекса.

Точность ассоциации гомологичных хромосом у *S. cerevisiae* контролируется геном *hop2*. У мутантов по *hop2* при мейозе синаптонемные комплексы образуются в нормальных количествах, но каждый отдельно взятый комплекс содержит негомологичные хромосомы. Таким образом, образование синаптонемных комплексов как таковых не зависит от гомологии (и потому не может быть основано на сколько-нибудь протяженном сравнении последовательностей ДНК). В норме роль *hop2* состоит в удержании негомологичных хромосом от взаимодействия.

В синаптонемных комплексах неверно спаренных хромосом у мутантов *hop2*⁻ двухцепочечные разрывы образуются, но не репарируются. Получается, что в то время как образование синаптонемного комплекса требует двухцепочечных разрывов, оно не требует сколько-нибудь экстенсивной реакции этих разрывов с участием гомологичной ДНК.

Неясно, что происходит в пахитене до того, как детектируются ДНК-рекомбинанты. Возможно, этот период отведен для поздних этапов рекомбинации (распространение зоны обмена цепями, синтез ДНК, разделение структуры Холлидея).

На следующей стадии мейоза — диплотене — синаптонемный комплекс распадается; хиазмы тогда становятся видны как точки, по которым хромосомы связаны. Это всегда расценивалось как индикатор обмена генетической информацией, однако молекулярная природа хиазм неизвестна. Возможно, они представляют собой остаточное явление прошедшего обмена или же связь между гомологичными хромосомами сохраняется там, где процесс обмена генетической информацией еще не получил разделения. Позднее в мейозе хиазмы перемещаются к концам хромосом. Такая подвижность говорит о том, что хиазмы отображают, скорее, остаток рекомбинации, чем действительный промежуточный продукт.

Рекомбинационные события происходят в разрозненных сайтах мейотических хромосом, но пока мы не можем привязать их к дискретным визуализуемым структурам — рекомбинационным узелкам и хиазмам. Докопаться до молекулярных основ образования разрозненных структур помогает идентификация белков, участвующих в рекомбинации у дрожжей, с «привязкой» их к дискретным сайтам. В числе таких белков — MSH4 (родственник бактериальных белков, участвующих в репарации неверного спаривания), а также Dmc1 и Rad51 (которые гомологичны белку RecA *E. coli*). Точные функции этих белков в рекомбинации не установлены.

Рекомбинационные процессы подвержены общему контролю. Лишь небольшое число единиц взаимодействия в реальности дозревает как кроссоверы, и распределены они таким образом, что, в общем случае, каждая пара гомологов обретает по одному или по два кроссовера; вместе с тем вероятность отсутствия кроссоверов у пары гомологов очень низка ($<0,1\%$). Возможно, это результат работы единой системы **кроссоверного контроля**, потому что неслучайность кроссоверов в целом нарушена у некоторых мутантов. Более того, наличие рекомбинации необходимо для течения мейоза, и существует «контрольно-пропускная» система, блокирующая мейоз, если рекомбинации не произошло. Блок снимается, если рекомбинация успешно завершена; такая система предохраняет клетку от попыток сегрегировать хромосомы прежде, чем совершится рекомбинация.

19.8 Бактериальная система RecBCD стимулируется последовательностями *chi*

Ключевые понятия:

- Комплекс RecBCD обладает нуклеазной и хеликазной активностями.
- Он связывается с ДНК ниже последовательности *chi*, расплетает дуплекс и разрушает одну из цепей в направлении 3′–5′ по мере продвижения к сайту *chi*.
- В сайте *chi* комплекс теряет субъединицу RecB и начинает проявлять нуклеазную активность.

Природа явлений, связанных с обменом цепями между молекулами ДНК, была впервые охарактеризована для бактериальных систем. Здесь реакция распознавания составляет неотъемлемую часть механизма рекомбинации и задействует ограниченные области молекул ДНК, а не хромосомы целиком. При этом общий порядок молекулярных событий у бактерий и эукариот похож: одиночная цепь из разорванной молекулы взаимодействует с дуплексом-партнером; область молекулярного спаривания распространяется; наконец, эндонуклеаза разделяет участников. Ферменты каждого этапа известны, но возможно, что они — лишь часть компонентов, необходимых для рекомбинации.

Бактериальные ферменты, задействованные в рекомбинации, были идентифицированы по наличию в их генах мутаций *rec*⁻. По фенотипу мутанты *rec*⁻ отличаются неспособностью к общей (гомологичной) рекомбинации. Всего было идентифицировано, по разным оценкам, от десяти до двадцати таких локусов.

Бактериям, как правило, не свойственна перестановка больших массивов двухцепочечной ДНК, но у прокариот могут быть разные пути инициации рекомбинации. В некоторых случаях используется ДНК со свободными одноцепочечными 3′-концами: ДНК может «поступать» в одноцепочечной форме (как при конъюгации; см. разд. 16.10 «Перенос Т-ДНК происходит аналогично переносу ДНК при конъюгации бактерий»); одноцепочечные бреши могут возникать при облучении; одноцепочечные «хвосты» могут также порождаться фаговыми геномами, проходящими репликацию по типу «катящегося кольца». Все же в ситуации, когда участвуют две дуплексных молекулы (как при рекомбинации в мейозе у эукариот), одноцепочечные участки и 3′-концы необходимо генерировать.

Один из механизмов генерации подходящих концов был открыт в связи с существованием определенных «горячих точек», стимулирующих рекомбинацию. Эти «горячие точки», открытые у фага лямбда в форме мутантов *chi*, содержат однонуклеотидные замены, создающие последовательности, стимулирующие рекомбинацию. Эти сайты «подвели нас» к белкам-участникам рекомбинации.

Общей для этих сайтов является константная несимметричная последовательность из 8 пн:

5′ GCTGGTGG 3′
3′ CGACCACC 5′

Последовательность *chi* естественным образом встречается в ДНК *E. coli* примерно один раз на каждые 5–10 тпн. Ее отсутствие в ДНК фага лямбда дикого типа, а также в других генетических элементах, показывает, что для рекомбинации она не существенна.

Последовательность *chi* стимулирует рекомбинацию на интервале до 10 тпн от себя. Сайт *chi* может быть активирован двухцепочечным разрывом, сделанным на расстоянии в несколько тпн по одну его конкретную сторону (вправо от последовательности, приведенной выше). Эта зависимость от ориентации наводит на предположение, что аппарат рекомбинации должен ассоциировать с ДНК на краю разрыва, и может затем двигаться вдоль дуплекса лишь в одном направлении.

Сайты *chi* являются целевыми для фермента, кодируемого генами *recBCD*. Этот комплекс проявляет несколько активностей. Во-первых, это мощная нуклеаза, разрушающая ДНК, исходно идентифицированная как активность экзонуклеазы V. Во-вторых, работая как хеликаза, он может расплести дуплекс ДНК в присутствии связывающего одноцепочечную ДНК белка SSB. В-третьих, он обладает АТФазной активностью. Его роль в рекомбинации, возможно, состоит в обеспечении одноцепочечного участка со свободным 3′-концом.

На РИС. 19.14 показано, как скоординированы эти реакции на субстрате ДНК, содержащей сайт *chi*. RecBCD связывает двухцепочечный конец ДНК. Две из его субъединиц обладают хеликазной активностью: RecD функционирует с 5′–3′-полярностью, а RecB функционирует с 3′–5′-полярностью. Сначала транслокацию вдоль ДНК и расплетание двойной спирали ведет субъединица RecD. По мере продвижения RecBCD разрушает высвобождающуюся одиночную цепь с 3′-концом. Когда комплекс достигает сайта *chi*, он узнает верхнюю цепь сайта *chi* в ее одиночной форме. Это заставляет фермент прервать движение. После этого он разрезает верхнюю цепь ДНК на расстоянии от четырех до шести нуклеотидов вправо от *chi*. Узнавание сайта *chi* приводит к диссоциации или инактивации субъединицы RecD, и на этом фермент лишается нуклеазной активности. Однако он продолжает работать хеликазой — теперь используя как средство передвижения лишь субъединицу RecB — примерно в половину прежней скорости. Итог всему — образование одноцепочечной ДНК с последовательностью *chi* на свободном 3′-конце. Это — идеальный субстрат для рекомбинации.

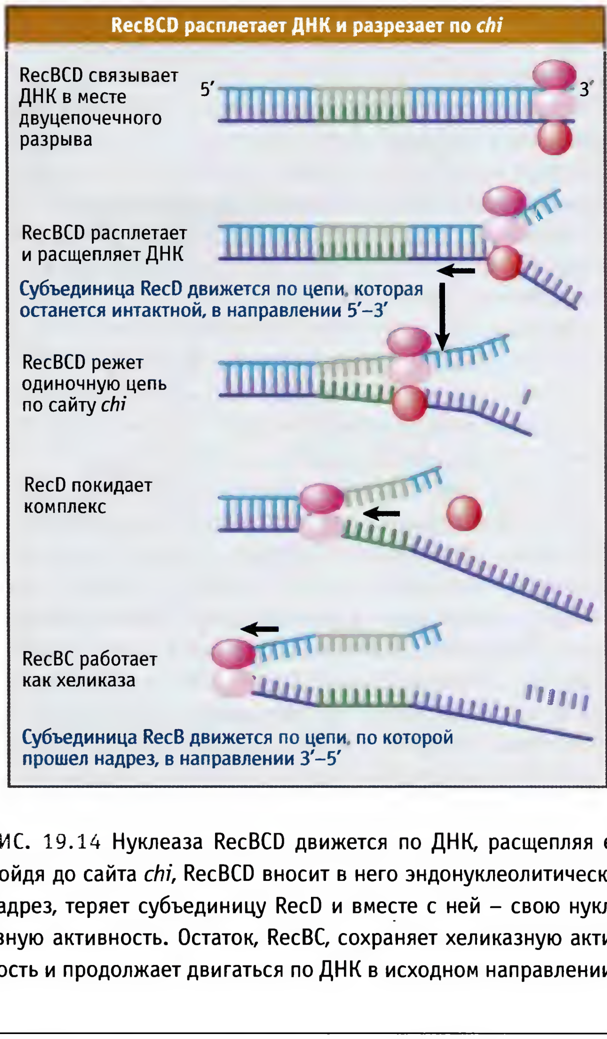


РИС. 19.14 Нуклеаза RecBCD движется по ДНК, расщепляя ее. Дойдя до сайта *chi*, RecBCD вносит в него эндонуклеолитический надрез, теряет субъединицу RecD и вместе с ней – свою нуклеазную активность. Остаток, RecBC, сохраняет хеликазную активность и продолжает двигаться по ДНК в исходном направлении

19.9 Белки переноса цепи катализируют ассимиляцию одноцепочечной ДНК

Ключевые понятия:

- **Белок RecA, образуя филаменты с ДНК, катализирует вторжение свободного 3′-конца одноцепочечной ДНК в дуплекс и замещение ею цепи-двойника.**

Белок RecA *E. coli* был первым примером белка-переносчика цепи ДНК. Принцип, заложенный в этот термин, объединяет несколько белков бактерий и архебактерий, белок *S. cerevisiae* Rad51 и белок Dmc1 высших эукариот. Анализ мутантов по *rad51* у дрожжей показывает, что этот класс белков играет центральную роль в рекомбинации. У таких мутантов накапливаются двухцепочечные разрывы и нормальные синаптонемные комплексы не образуются. Это подкрепляет мнение, что обмен цепями между дуплексами ДНК причастен к образованию синаптонемного комплекса, и указывает на возможность родства между процессами синапсиса эукариотических хромосом и ассимиляции одиночной цепи ДНК у бактерий.

Бактериальный белок recA обладает двумя совершенно разными типами активности. Он может стимулировать протеазную активность при SOS-ответе (см. разд. 20.10 «RecA запускает SOS-систему») и может промотировать спаривание оснований между одноцепочечной ДНК и ее комплементом в молекуле-дуплексе. Обе активности приводятся в действие одноцепочечной ДНК в присутствии АТФ.

ДНК-обрабатывающая активность RecA приспосабливает одиночную цепь к замещению своего гомолога в дуплексе в ходе реакции, получившей название **поглощения**, или **ассимиляции**, **одиночной цепи**. Реакция замещения может произойти между молекулами ДНК в нескольких конфигурациях при соблюдении трех условий:

- Одна из молекул ДНК должна иметь одноцепочечный участок.
- Одна из молекул должна иметь свободный 3′-конец.
- Одноцепочечный участок и 3′-конец должны находиться внутри участка комплементарности этих молекул.

Реакцию иллюстрирует рис. 19.15. Когда линейная одиночная цепь вторгается в дуплекс, она вытесняет исходного партнера своего комплемента. За реакцией проще всего следить, сделав донора либо реципиента кольцевой молекулой. Реакция развивается в направлении 5′–3′ вдоль цепи, чей партнер подлежит вытеснению и замещению, и заключается в перестановке цепей, (как минимум), одна из которых имеет свободный 3′-конец.

Ассимиляция одиночной цепи потенциально родственна инициации рекомбинации. Обе модели объединяет неизбежность промежуточного продукта, в котором одна или обе одиночные цепи кроссингуют с одного дуплекса на другой (см. рис. 19.6 и рис. 19.9). Этот этап катализирует RecA. В бактериальном контексте, RecA действует на субстраты, порожденные RecBCD. Опосредованное RecBCD расплетание и разрез могут быть использованы для создания концов, которые, в свою очередь, инициируют образование гетеродуплексных соединений. RecA может взять одиночную цепь с 3′-концом, высвободившимся при надрезе RecBCD по сайту *chi*, и заставить взаимодействовать с гомологичной последовательностью дуплекса, создавая таким образом соединенную молекулу.

Все белки семейства переносчиков цепи у бактерий и архебактерий могут агрегировать с одноцепочечной

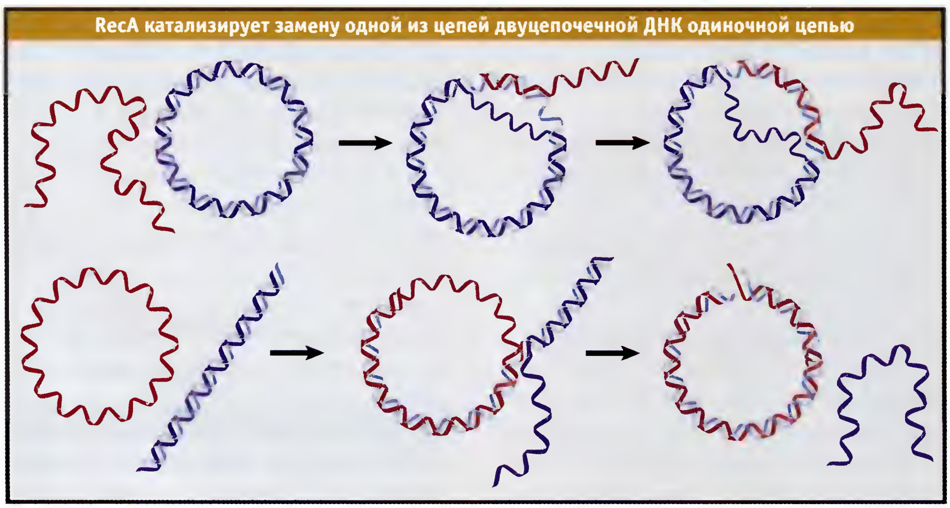


Рис. **Белок RecA обеспечивает встраивание чужеродной одиночной цепи вместо одной из цепей исходного дуплекса; для этого у одной из конкурирующих цепей должен быть свободный конец**

или дуплексной ДНК, образуя длинные филаменты. На виток филамента (который имеет спиралевидную структуру с глубоким желобком, содержащим ДНК) приходится шесть мономеров RecA. Стехиометрия связывания равняется трем нуклеотидам (или парам оснований) на один мономер RecA. ДНК удерживается в форме (в полтора раза растянутой по отношению к дуплексу ДНК В), на виток которой приходится 18,6 нуклеотидов (или пар оснований). Когда связан дуплекс ДНК, он контактирует с RecA своей малой бороздкой, оставляя большую бороздку доступной для возможной реакции с другой молекулой ДНК.

Взаимодействие между двумя молекулами ДНК происходит внутри таких филаментов. При поглощении одиночной цепи дуплексом RecA прежде всего увязывает одиночную цепь в филамент. Затем заключает туда же дуплекс, возможно, с образованием какой-то разновидности трехцепочечной структуры. В этой системе синапсис (сближение) молекул предшествует физическому обмену материалом, потому как реакция сближения молекул может иметь место даже в отсутствие свободных концов, когда обмен цепями невозможен. Обмен цепями, при участии свободного 3′-конца, происходит внутри филамента, и RecA остается связан с цепью, бывшей исходно одиночной, так что в конце реакции RecA оказывается связан с дуплексной молекулой.

Все белки семейства могут промотировать основной процесс обмена цепями, не тратя энергию. RecA, однако, интенсифицирует эту свою активность, пользуясь гидролизом АТФ. По ходу реакции АТФ гидролизуется в большом количестве. АТФ может влиять на конформацию RecA посредством аллостерического эффекта. В контакте с АТФ ДНК-связывающий сайт RecA проявляет высокое сродство к ДНК; это нужно для связывания ДНК и для реакции спаривания. Гидролиз АТФ переводит ДНК-связывающий сайт RecA в конформацию с низким сродством к ДНК, которая необходима для высвобождения гетеродуплекса.

Катализируемую RecA реакцию между одно- и двухцепочечной ДНК можно разделить на три фазы:

- медленная «пресинаптическая» фаза, в которой RecA полимеризуется на одноцепочечной ДНК;
- быстрое спаривание одиночной цепи с комплементом в дуплексе с образованием гетеродуплексного соединения; и
- медленное замещение одной из цепей дуплекса с образованием протяженного участка гетеродуплексной ДНК.

Присутствие белка SSB стимулирует реакцию, обеспечивая отсутствие у субстрата вторичной структуры. Все же непонятно, как SSB и RecA вдвоем могут действовать на один и тот же сегмент ДНК. Так же как для SSB, расход RecA пропорционален течению реакции, и, наверное, он так же, как SSB, проявляет кооперативность при связывании ДНК для образования филаментоподобной структуры.

Когда одноцепочечная молекула ДНК взаимодействует с дуплексом, то он оказывается расплетен на участке рекомбинантного соединения. Первый участок гетеродуплекса может даже не укладываться в общепринятый двуспиральный формат, но все же состоять из двух цепей, ассоциированных бок о бок. Участок такого типа называют **паранемическим соединением** в противоположность классическому **плектонемическому** переплетению цепей в двойной спирали. Паранемическое соединение нестабильно; дальнейший ход реакции обусловлен его переходом в форму двойной спирали. Этот переход эквивалентен сбросу отрицательных сверхвитков, и для него может быть необходим фермент, который решает проблему расплетания/сплетания с помощью внесения временных разрывов, позволяющих цепям вращаться друг относительно друга.

Все реакции, которые мы обсуждали до сих пор, составляют лишь часть потенциального события рекомбинации: вторжение в один дуплекс одиночной цепи. Две дуплексных молекулы могут взаимодействовать друг с другом при поддержке RecA при условии, что одна из них содержит одноцепочечный участок протяженностью как минимум в 50 нуклеотидов. Одноцепочечный участок может быть в форме «хвоста» у линейной молекулы или в форме бреши в кольцевой молекуле.

Реакция между частично дуплексной и полностью дуплексной молекулами ведет к обмену цепями. Пример изображен на РИС. 19.16. Ассимиляция начинается на одном конце линейной молекулы, где вторгающаяся одиночная цепь замещает своего гомолога в дуплексе по обычной схеме. Когда реакция достигает участка, двухцепочечного у обеих молекул, вторгающаяся в другой дуплекс цепь покидает своего изначального партнера, который затем спаривается с другой (вытесняемой)-цепью.

На этом этапе молекула имеет структуру, неотличимую от рекомбинантного соединения с рис. 19.8. Реакция, поддерживаемая RecA *in vitro*, может порождать структуры Холлидея, и это значит, что фермент может опосредовать реципрокный перенос цепей. О геометрии связанных RecA четырехцепочечных интермедиатов мало что известно; предположительно две дуплексные молекулы могут располагаться бок о бок в манере, соответствующей требованиям реакции обмена.

Характеризация биохимических реакций *in vitro* оставляет открытыми множество возможностей функционирования белков переноса цепи *in vivo*. Их участие инициируется наличием одноцепочечного 3′-конца. У бактерий он чаще всего образуется вследствие процессинга двухцепочечного разрыва белком RecBCD. Одним из самых частых обстоятельств подключения этой опции является репликативная вилка, «застрявшая» на месте повреждения ДНК (см. разд. 20.9 «Рекомбинация служит важным механизмом исправления ошибок репликации»). Перенос ДНК при конъюгации, когда RecA необходим для рекомбинации с хромосомой хозяина, уже ближе к рекомбинации в традиционном смысле. У дрожжей двухцепочечные разрывы могут быть результатом как повреждения ДНК, так и нормального хода

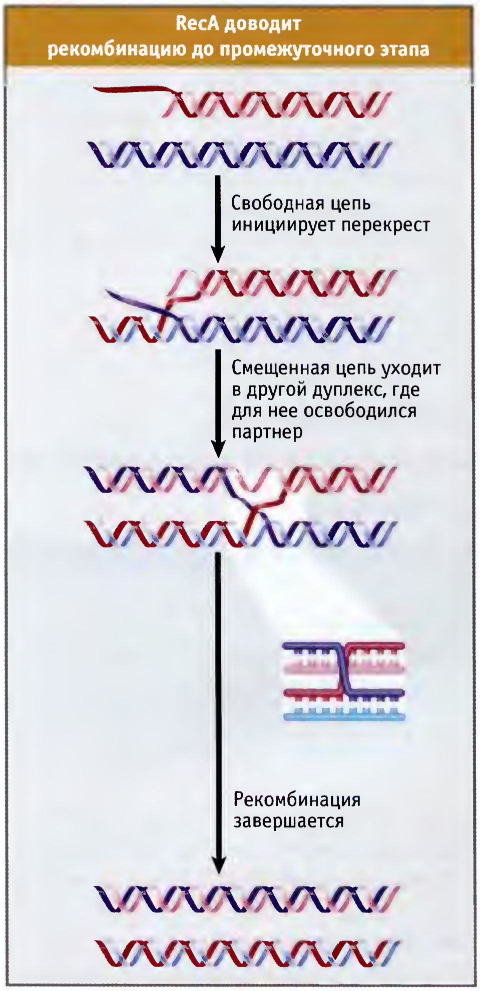


РИС. 19.16 Белок RecA катализирует обмен цепями между частичным и полным дуплексами ДНК. Получается соединенная молекула, эквивалентная структуре Холлидея

рекомбинации. В обоих случаях за обработкой разрыва с образованием одноцепочечного 3′-конца следует загрузка одиночной цепи в филамент с Rad51 для поиска идентичных дуплексных последовательностей. Это может использоваться при репарации, равно как при рекомбинации.

19.10 Система Ruv разделяет структуры Холлидея

Ключевые понятия:

- Комплекс Ruv обрабатывает стыки рекомбинирующих молекул.
- Белок RuvA опознает структуру стыка, а RuvB является хеликазой, катализирующей миграцию ветви.
- Белок RuvC разрезает стыки с образованием промежуточных продуктов рекомбинации.

Один из важнейших этапов рекомбинации — разрезание структуры Холлидея, определяющее исход всего процесса как реципрокную рекомбинацию или как возврат в исходное (не считая короткого участка гибридной ДНК) состояние (см. рис. 19.6 и рис. 19.8). Миграция ветви от сайта обмена (см. рис. 19.7) детерминирует длину участка гибридной ДНК (с рекомбинацией или без нее). Белки, участвующие в стабилизации и разрезании структур Холлидея у *E. coli*, были идентифицированы как продукты генов *ruv*. Белки RuvA и RuvB интенсифицируют образование гетеродуплексных структур. RuvA распознает структуру Холлидея, связывается со всеми четырьмя цепями ДНК в точке перекреста, и два его тетрамера образуют с ДНК сандвич-структуру. RuvB — гексамерная хеликаза с АТФазной активностью, он служит мотором для миграции ветви. Гексамерные кольца RuvB охватывают каждый дуплекс ДНК выше точки перехода (относительно направления миграции ветви). Комплекс схематично изображен на РИС. 19.17.

Комплекс RuvAB может заставлять ветвь мигрировать со скоростью от 10 до 20 пн/с. Похожую активность обеспечивает другая хеликаза, RecG. В продолжение своего действия RuvAB смещает RuvA с ДНК. Активности RuvAB и RecG обе могут действовать на структуры Холлидея, и если обе нарушены вследствие мутаций, рекомбинантная активность в клетке *E. coli* безнадежно дефектна.

Третий ген, *ruvC*, кодирует эндонуклеазу, специфично узнающую структуры Холлидея. Она может разрезать соединения *in vitro*, осуществляя, таким образом, разделение промежуточных продуктов рекомбинации. Общая тетрануклеотидная последовательность обеспечивает для RuvC горячую точку разделения структуры Холлидея. Тетрануклеотид ATTG асимметричен, и, таким образом, может направлять разделение структуры относительно того, какую пару цепей следует разрезать. Этот выбор определяет либо образование заплаточных рекомбинантов (то есть отсутствие полноценной рекомбинации), либо образование сплайс-рекомбинантов (рекомбинация между фланкирующими маркерами). Анализ кристалли-

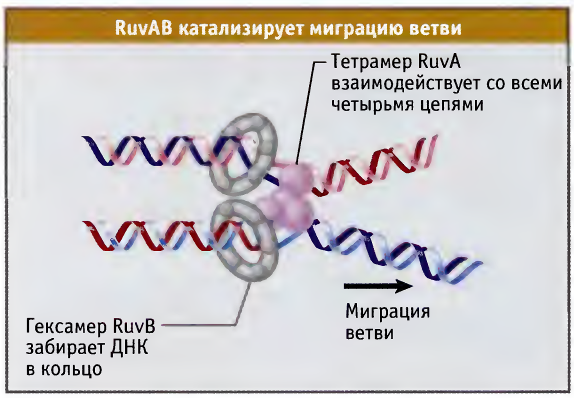
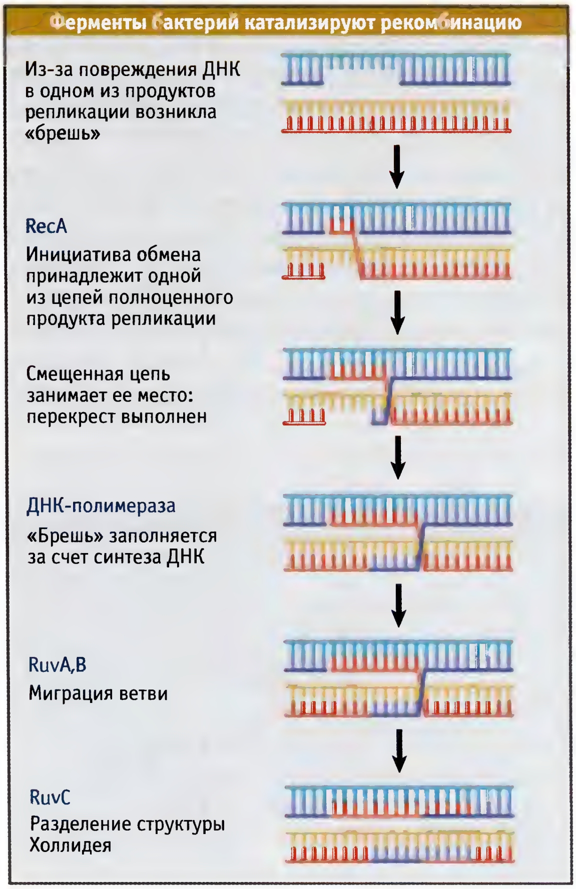


РИС. 19.17 RuvAB представляет собой асимметричный комплекс, обеспечивающий перемещение точки ветвления (миграцию ветви) в структуре Холлидея

ческих структур RuvC и других разделяющих ферментов показал, что их структуры мало похожи, несмотря на общность функции.

Учитывая все вышесказанное, можно предположить, что на рекомбинацию работает разделяющий белковый комплекс, в который входят и ферменты-катализаторы миграции ветви, и собственно разделяющие белки. Возможно, похожий комплекс присутствует в клетках млекопитающих.

Теперь мы в состоянии объяснить стадии рекомбинации у *E. coli* функциями конкретных белков. На рис. 19.18 показаны этапы использования рекомбинации для репарации пробела в одном дуплексе с привлечением материала из другого (гомологичного) дуплекса. Главное препятствие для переноса этих знаний на рекомбинацию у эукариот состоит в том, что у бактерий в общем случае целая хромосома рекомбинирует с фрагментом ДНК. Такая рекомбинация носит характер репарации, стимулируемой повреждением ДНК, и это не совсем эквивалентно рекомбинации между геномами при мейозе. И все же эукариоты манипулируют своей ДНК с использованием похожих молекулярных активностей.



Ферменты бактерий способны вести все этапы репарационной рекомбинации, если их снабдить подходящими молекулами ДНК в качестве субстрата

Другая система разделяющих ферментов, или резольваз, была описана у дрожжей и млекопитающих. У мутантов *S. cerevisiae* по *mus81* нарушена рекомбинация. Белок Mus81 — компонент эндонуклеазы, которая разделяет структуры Холлидея на отдельные дуплексы. Эта резольваза важна как при мейозе, так и для рестарта «застрявших» репликативных вилок (см. разд. 20.9 «Рекомбинация служит важным механизмом исправления ошибок репликации»).

19.11 Конверсия гена как механизм межаллельной рекомбинации

Ключевые понятия:

- **Гетеродуплекс ДНК, возникший в ходе рекомбинации, содержит некомплементарные основания, если рекомбинирующие последовательности не идентичны.**
- **Системы репарации могут устранять несоответствия, изменяя последовательность одной из цепей так, чтобы она стала комплементарна другой**

Характерные свойства рекомбинации между аллелями объясняются участием в ней гетеродуплексной ДНК; действительно, феномен межаллельной рекомбинации побуждал к разработке гетеродуплексной модели. Когда рекомбинация между аллелями была открыта, естественно было предположить, что механизм ее тот же, что и у реципрокной рекомбинации между удаленными друг от друга локусами. Иными словами, внутри локуса происходит индивидуальный акт разрыва и воссоединения, порождающий реципрокную пару рекомбинантных хромосом. Для небольшого участка одного-единственного гена само образование гетеродуплексной ДНК обычно составляет суть рекомбинации.

Индивидуальные факты рекомбинации можно изучать на грибах-аскомицетах, поскольку у них продукты одиночного мейоза удерживаются вместе в большой клетке — **аске** (менее употребительное название — **тетрада**). Более того, у некоторых из этих грибов четыре гаплоидных ядра — продукта мейоза — расположены в линейном порядке (последующее митотическое деление этих четырех ядер дает линейную серию из восьми гаплоидных ядер). На показано, что каждое из этих ядер представляет генетический характер одной из восьми цепей четырех хромосом, сформировавшихся в результате мейоза.

При мейозе гетерозиготной клетки образуется по четыре копии каждого аллеля. Именно это наблюдается в большинстве спор. Однако в некоторых спорах соотношение другое. Это объясняется образованием и коррекцией гетеродуплексной ДНК на участке, по которому аллели отличаются. Рисунок иллюстрирует событие рекомбинации, при котором на одной из четырех мейотических хромосом оказывается отрезок

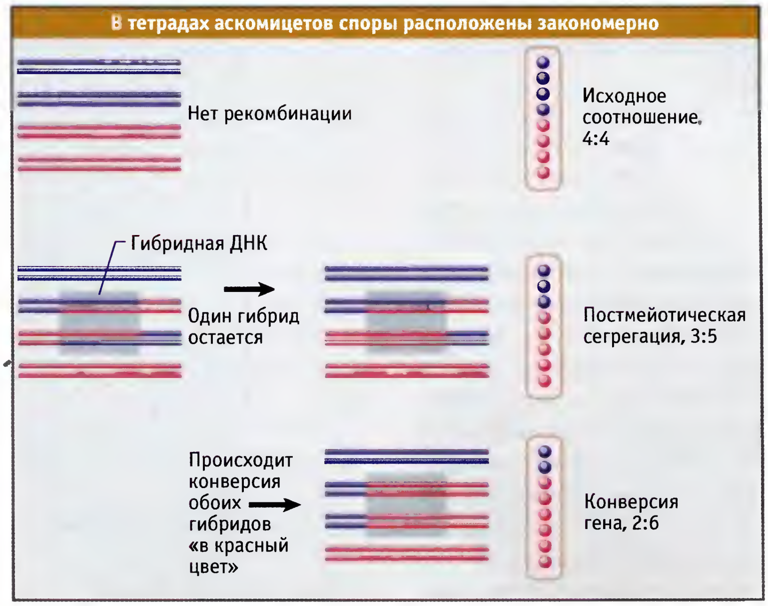


РИС. 19.19 При спорообразовании у аскомицетов можно проследить мейотическую судьбу каждой отдельно взятой цепи ДНК

гибридной ДНК, возникший, возможно, как результат рекомбинации, инициированной двухцепочечным разрывом.

Предположим, два аллеля различаются по одной точечной мутации. Когда в результате обмена цепями ДНК образуется гетеродуплекс, его цепи в точке мутации не будут комплементарны. При этом каждая цепь будет нести свою генетическую информацию. Если в их последовательности не вносится изменений, с наступлением репликации цепи стабилизируют свою информацию в составе дочернего дуплекса. Такой вариант развития событий называют **постмейотической сегрегацией**, что означает «разделение (цепей ДНК) после мейоза». Его значение состоит в том, что он напрямую демонстрирует существование гетеродуплексной ДНК в рекомбинирующих аллелях.

Другой эффект проявляется при исследовании рекомбинации между аллелями: пропорция их частот отличается от исходного соотношения 4:4. Этот эффект называют **конверсией гена**. Он заключается в нереципрокном переносе информации с одной хроматиды на другую.

Конверсия гена проистекает из обмена цепями между молекулами ДНК, и изменение последовательности может иметь одну из двух возможных причин молекулярного уровня:

- Как обозначено на рис. 19.9, в модели с двухцепочечным разрывом один дуплекс ДНК служит донором генетической информации, которая напрямую замещает сответствующие последовательности в дуплексе-реципиенте в ходе процессов образования «бреши», обмена цепями и заполнения «бреши».
- Как часть обменного процесса гетеродуплекс ДНК образуется, когда одиночная цепь одного дуплекса взаимодействует со своим комплементом в другом дуплексе. Системы репарации узнают неспаренные основания в гетеродуплексной ДНК и, чтобы восстановить комплементарность, могут вырезать и заменить одну из цепей. Этот процесс преобразует цепь ДНК, представляющую один аллель, в последовательность другого аллеля.

Конверсия гена происходит вне зависимости от кроссинговера, но коррелирует с ним. Значительная часть аберрантных асков демонстрирует генетическую рекомбинацию между двумя маркерами по обе стороны сайта межаллельной конверсии гена. Это соответствует ситуации, в которой аберрантные соотношения — результат инициации рекомбинативного процесса (в соответствии с рис. 19.6), но с приблизительно одинаковой вероятностью разводки структуры с рекомбинацией или без нее (как обозначено на рис. 19.8). Из этого логически следует, что хромосомы гриба инициируют кроссинговер примерно вдвое чаще, чем можно ожидать исходя из измерений частоты рекомбинации между отдаленными генами.

Изучение рекомбинации на молекулярном уровне выявляет разнообразные тенденции. Направленность конверсии гена может быть равновероятной, или же аллель-специфичные эффекты могут создавать приоритет одной из сторон. Градиенты рекомбинации могут быть разобщены с горячими точками. Сейчас нам известно, что горячие точки представляют собой сайты, по которым инициируются двухцепочечные разрывы, а градиент коррелирует со степенью разрастания пробела

с длинными одноцепочечными концами по краям (см. разд. 19.6 «Синаптонемный комплекс образуется после внесения двухцепочечных разрывов»).

Некоторую информацию о степени конверсии гена можно почерпнуть, анализируя последовательности кластерных генов. Обычно продукты рекомбинации разделяются и становятся бесполезны для анализа рекомбинации на уровне последовательности ДНК. Если же хромосома несет два (неаллельных) гена, родственных друг другу, они могут рекомбинировать посредством «неравного кроссинговера» (см. разд. 6.7 «Неравный кроссинговер приводит к реорганизации целых генетических кластеров»). Сейчас нам нужно лишь отметить, что гетеродуплекс может образовываться между двумя неаллельными генами. Конверсия гена фактически преобразует один из неаллельных генов в последовательность другого.

Присутствие более чем одной копии гена на одной и той же хромосоме — «отпечаток» такого события. Например, если образование гетеродуплекса и конверсия гена охватили лишь часть гена, то эта часть имеет последовательность, идентичную (или очень близкородственную) другому гену, в то время как для остальной части их последовательности отличаются. Исследования показывают, что конверсия гена может охватить значительный промежуток, вплоть до нескольких тысяч нуклеотидов.

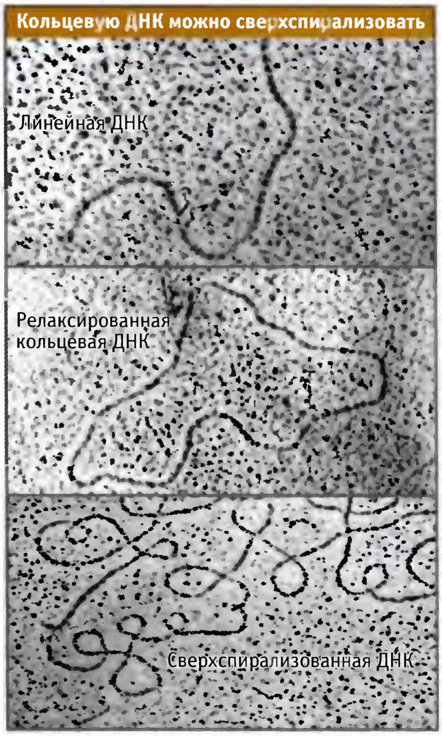


РИС. 19.20 Свободная линейная ДНК выглядит расправленной. Кольцевая ДНК выглядит расправленной, если она релаксирована (то есть не содержит сверхвитков). Сверхспирализованная ДНК образует петли и выглядит конденсированной. Фотография любезно предоставлена Nirupam Roy Choudhury, International Centre for Genetic Engineering and Biotechnology (ICGEB)

19.12 Сверхспирализация затрагивает структуру ДНК

Ключевые понятия:

- Сверхспирализация происходит только в «закрытой» ДНК, не имеющей свободных концов.
- «Закрытой» является ДНК, замкнутая в кольцо, или в линейной форме с концами, закрепленными белковой структурой.
- Всякая «закрытая» молекула ДНК характеризуется порядком зацепления, представляющим собой сумму числа витков двойной спирали и числа сверхвитков ($Lk = Tw + Wr$).
- Посредством перераспределения витков изменение структуры двойной спирали может компенсировать изменение ее скручивания в пространстве.
- Порядок зацепления может быть изменен только с помощью разрыва химических связей в ДНК и образования новых.

Закрученность цепей ДНК друг относительно друга в структуре двойной спирали делает возможным менять эту структуру, воздействуя на ее конформацию в пространстве. Если неподвижно зафиксировать оба конца молекулы ДНК, двойную спираль такой «закрытой» молекулы можно вращать относительно себя в пространстве. Это называют **сверхспирализацией**. Результат можно представить как резиновую ленту, закручиваемую относительно себя. Простейшим примером ДНК без свободных концов является кольцевая молекула. Эффект сверхспирализации можно увидеть, сравнивая несверхспирализованную кольцевую ДНК, плоско лежащую на фото с РИС. 19.20, со сверхспирализованной кольцевой молекулой.

Последствия сверхспирализации зависят от того, в какую сторону закручена ДНК относительно себя. Если в ту же, что и две цепи внутри двойной спирали (по часовой стрелке), то это *положительная сверхспирализация* (*positive supercoiling*). Она заставляет цепи ДНК закручиваться друг относительно друга более плотно, с большим числом пар оснований на виток. Вращение ДНК относительно себя в противоположной ориентации порождает *отрицательную сверхспирализацию* (*negative supercoiling*). При ней цепи ДНК становятся закручены менее плотно, с меньшим, чем в исходном (релаксированном) состоянии, числом пар оснований на виток. Отрицательную сверхспирализацию можно представлять себе как создание в ДНК напряжений, которые сбрасываются посредством расплетания двойной спирали. Крайним проявлением эффекта негативной сверхспирализации является участок с разделенными цепями ДНК — формально, с нулем пар оснований на виток.

Топологическое манипулирование ДНК можно считать центральным аспектом всех ее функциональных активностей (рекомбинации, репликации, транскрипции), а также организации структуры высшего

порядка. Все синтетические активности с участием двухцепочечной ДНК требуют разделения цепей. Но, оказывается, цепи не просто лежат бок о бок; они переплетаются. Поэтому их разделение требует вращения друг относительно друга в пространстве. Некоторые возможности для реакции расплетания проиллюстрированы на РИС. 19.21.

Можно представить себе ДНК со свободным концом, который позволил бы цепям вращаться вокруг оси двойной спирали и расплетаться. Однако неизбежное при этом «болтание» цепей, учитывая длину двойной спирали и внутриклеточную пространственную стесненность, не выглядит реалистично.

Если поместить на свободный конец ДНК «устройство», контролирующее его вращение, это предотвратит раскручивание цепи, однако эффект должен быть передан на значительное расстояние, опять же с использованием ротации несообразно длинного участка.

Рассмотрим последствия разделения цепей в молекуле, чьи концы не свободны для вращения. Если две переплетенные цепи взять за концы и растянуть врозь, результатом будет уплотнение их закручивания в дальней части молекулы. Проблему можно решить внесением временного разрыва (ника) в одну из цепей. Внутренний свободный конец позволяет надрезанной цепи вращаться вокруг интактной цепи, после чего ник может быть зашит. Каждое повторение реакции надрезания и сшивания высвобождает один сверхспиральный виток.

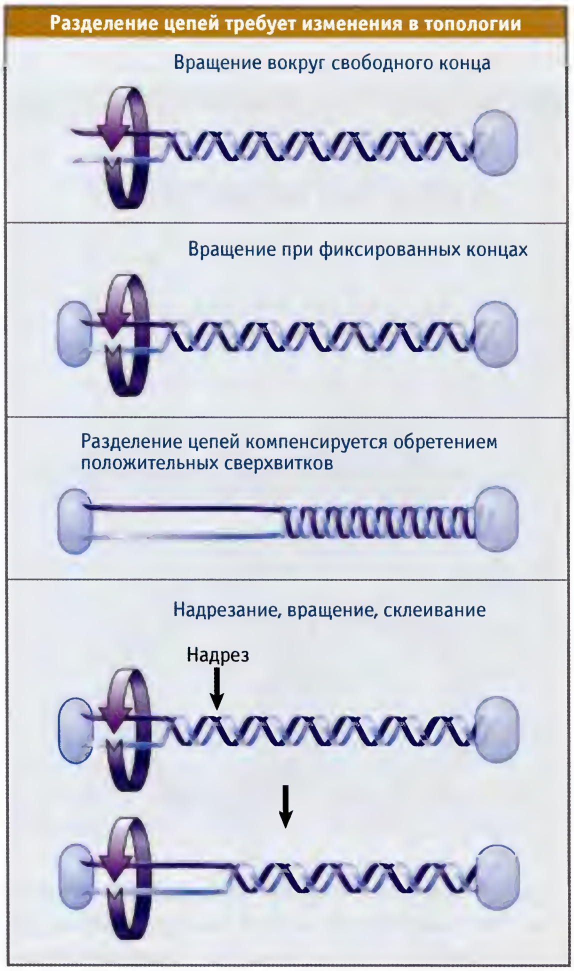


РИС. 19.21 Разделение цепей двойной спирали ДНК можно произвести несколькими способами

Закрытую молекулу ДНК характеризует ее **порядок зацепления** (L), число раз, которое одна ее цепь пересекает другую в пространстве. Закрытые молекулы ДНК с идентичными последовательностями могут иметь разные порядки зацепления, отражающие разные степени сверхспирализации. Молекулы ДНК, одинаковые во всем кроме порядков зацепления, называют **топологическими изомерами**.

У порядка зацепления есть две составляющих: число сверхвитков (W) и число витков двойной спирали (T).

Число витков двойной спирали (T) — это свойство самой двуспиральной структуры, отображающее вращение одной цепи относительно другой. Оно отображает общее число витков дуплекса и определяется числом пар оснований на виток. Для релаксированной закрытой кольцевой ДНК, лежащей плашмя на плоскости, число витков равняется общему числу пар оснований, деленному на число пар оснований в одном витке.

Число сверхвитков (W) отображает вращение оси дуплекса в пространстве. Оно соответствует интуитивному представлению о сверхспирализации, но количественно это не совсем одинаковые понятия. Для релаксированной молекулы W = 0, и порядок зацепления равен числу витков.

Часто интерес вызывает изменение порядка зацепления, ΔL, заданного равенством

$$\Delta L = \Delta W + \Delta T.$$

Это равенство означает, что любое изменение общего числа обращений одной цепи ДНК относительно другой может быть выражено как сумма изменений спирализации оси дуплекса в пространстве (ΔW) и скрученности самой двойной спирали (ΔT). В свободной молекуле ДНК W и T свободно регулируются, и любое ΔL (изменение порядка зацепления), по-видимому, выражается изменением W, то есть изменением сверхспирализованности.

Уменьшение порядка зацепления, то есть отрицательное значение ΔL, соответствует внесению некоей комбинации отрицательной сверхспирализации и/или расплетания. Увеличение порядка зацепленя, измеряемое как положительное значение ΔL, соответствует уменьшению отрицательной сверхспирализации/расплетания.

Изменение состояния любой ДНК можно описать специфической разностью зацепления (specific linking difference), $\sigma = \Delta L/L_0$, где L_0 — порядок зацепления для релаксированной ДНК. Если изменение порядка зацепления целиком вытекает из изменения W (то есть

$\Delta T = 0$), специфическая разность зацепления равняется плотности сверхспирализации. Фактически, σ как $\Delta L/L_0$ можно полагать соответствующей плотности сверхскручивания при условии постоянства структуры самой двойной спирали.

Говоря о порядке зацепления, важно понимать, что этот показатель инвариантен для любой индивидуальной закрытой молекулы ДНК. Порядок зацепления не может быть изменен какой-либо деформацией, кроме как с использованием разрыва и заклеивания. У кольцевой молекулы с конкретным порядком зацепления этот порядок может выражаться в разных сочетаниях T и W, сумма которых, однако, не может измениться до тех пор, пока ни одна из цепей не будет разорвана. (На самом деле, распределение значения L, фиксированного для молекулы ДНК в растворе, между T и W означает свободу каждой из этих составляющих принимать разные значения.)

Порядок зацепления связан с действительными, энзиматически обусловленными событиями, посредством которых в топологию ДНК вносятся изменения. Порядок зацепления конкретной закрытой молекулы может быть изменен лишь разрывом цепи или цепей, использованием свободного конца для вращения одной цепи вокруг другой и заклеиванием рваных концов. Изменение порядка зацепления после осуществления ферментом такой операции должно выражаться целым числом; его значение можно определить как характеристику реакции. Как выглядит это изменение в терминах ΔW и ΔT, надо рассматривать отдельно.

19.13 Топоизомеразы изменяют число сверхвитков, релаксируя или сверхспирализуя ДНК

Ключевые понятия:

- Топоизомеразы, изменяя порядок зацепления, разрушают в ДНК химические связи, меняют конформацию двойной спирали в пространстве и восстанавливают связи.
- Топоизомеразы первого типа, работая, рвут одну цепь ДНК; топоизомеразы второго типа действуют с внесением двухцепочечных разрывов

Изменения в топологии ДНК могут быть вызваны несколькими способами. На РИС. 19.22 приведены некоторые примеры. Чтобы началась репликация или транскрипция, две цепи ДНК должны быть расплетены. В случае репликации две цепи разделяются навсегда, и каждая образует новый дуплекс с дочерней цепью. В случае транскрипции движение РНК-полимеразы создает область положительной сверхспирализации впереди и область отрицательной сверхспирализации позади фермента. Система борется с этим постольку, поскольку положительные сверхвитки затрудняют движение фермента (см. разд. 11.15 «Сверхспирализация ДНК влияет на процесс транскрипции»). При репликации кольцевой молекулы ДНК кольцевые продукты могут оказаться сцеплены (катенированы) с одним из них, пропущенным сквозь другой. Они должны быть разделены, чтобы дочерние молекулы могли сегрегировать в дочерние клетки разделились. И еще один случай, в котором сверхспирализация играет важную роль, имеет место в ядре эукариотической клетки — это укладка тяжа ДНК в виде цепочки нуклеосом (см. разд. 29.6 «В нуклеосоме молекулярная периодичность ДНК изменяется»). Все перечисленные ситуации разрешаются действием топоизомераз.

ДНК-топоизомеразы — это ферменты, катализирующие изменения топологии ДНК. Они вносят временный разрыв в одну или обе цепи ДНК, сквозь него пропускают интактную цепь (цепи), после чего этот разрыв заклеивают. Концы, образующиеся при разрыве, никогда не бывают свободными: фермент образует с ними ковалентные сшивки и манипулирует ими внутри собственных границ. Топоизомеразы работают с ДНК безотносительно ее последовательности, хотя некоторые ферменты-участники сайт-специфичной рекомбинации функционируют тем же путем и также подпадают

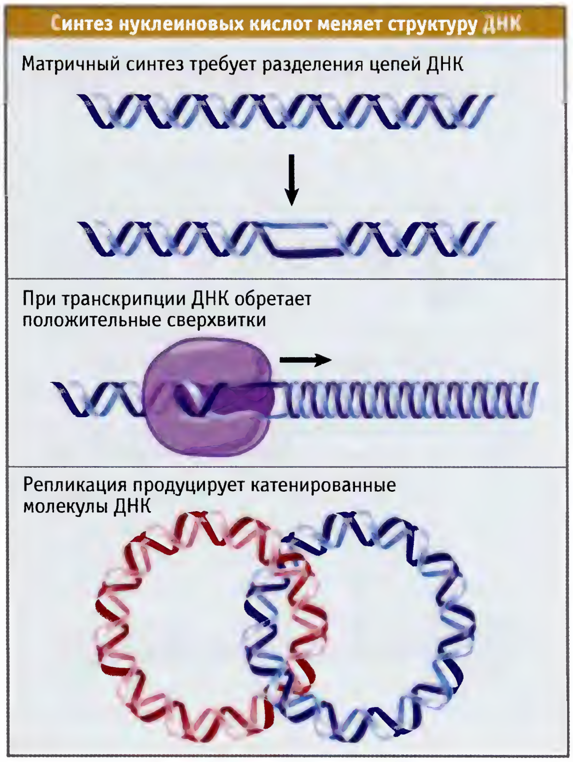


РИС. 19.22 Во время репликации и транскрипции топология ДНК меняется. Разделение цепей двойной спирали требует расплетания ее основных витков. Транскрипция создает положительные сверхвитки впереди РНК-полимеразы, а репликация кольцевой матрицы приводит к образованию сцепленных дочерних дуплексов (конкатамеров)

под определение топоизомераз (см. разд. 19.18 «Сайт-специфичная рекомбинация напоминает топоизомеразную активность»).

По природе используемых механизмов топоизомеразы подразделяют на два класса. **Топоизомеразы I типа** в своем действии производят разрыв одной цепи ДНК. **Топоизомеразы II типа** действуют с внесением двухцепочечного разрыва. В целом топоизомеразы различаются по типам вносимых ими топологических изменений. Некоторые топоизомеразы могут «расслаблять» (релаксировать, устранять) только отрицательные сверхвитки ДНК; другие — как отрицательные, так и положительные. Ферменты, способные вносить отрицательные сверхвитки, называют гиразами; способные же создавать положительные сверхвитки называют обратными гиразами.

У *E. coli* есть четыре фермента этой группы: топоизомеразы I, III и IV, а также ДНК-гираза. ДНК-топоизомеразы I и III — ферменты первого типа. Гираза и ДНК-топоизомераза IV — ферменты второго типа. Каждый из четырех ферментов важен для одной или более ситуаций, изображенных на рис. 19.22:

- Общий уровень отрицательной сверхспирализации в бактериальном нуклеоиде представляет собой результат равновесия между внесением сверхвитков гиразой и их релаксацией топоизомеразами I и IV. Это важнейший аспект структуры нуклеоида (см. разд. 28.4 «Бактериальный геном сверхспирализован»), затрагивающий инициацию транскрипции на определенных промоторах (см. разд. 11.15 «Степень сверхспирализации ДНК влияет на процесс транскрипции»).
- Те же ферменты участвуют в ликвидации проблем, создаваемых транскрипцией; гираза преобразует положительные сверхвитки, возникающие впереди РНК-полимеразы, в отрицательные, а топоизомеразы I и IV устраняют отрицательные сверхвитки позади РНК-полимеразы. Похожие, но более сложные, проблемы возникают при репликации, в их решении имеет место похожее распределение ролей между ферментами.
- В ходе репликации дочерние дуплексы могут оказаться закручены друг относительно друга на этапе, известном как прекатенация. Прекатенаны устраняются топоизомеразой IV, которая также декатенирует (расцепляет) геномы, оставшиеся как-либо катенированными по окончании репликации. Функции топоизомеразы III частично перекрываются с функциями топоизомеразы IV.

Ферменты эукариот следуют в своей работе тем же принципам, хотя детали могут отличаться. При этом по белковой последовательности и структуре они не похожи на ферменты прокариот. Большинство эукариот имеют единственную топоизомеразу первого типа, используемую и для продвижения репликативной вилки, и для релаксации сверхвитков, образующихся при транскрипции. Топоизомеразы второго типа требуются для разъединения хромосом после репликации. Особые разновидности топоизомераз участвуют в событиях рекомбинации и репарации.

19.14 Топоизомеразы разрывают цепи ДНК и снова их склеивают

Ключевые понятия:

- Топоизомеразы I типа вступают в ковалентную связь с одним из разорванных концов и после вращения одной цепи вокруг другой снова склеивают концы ДНК. Непрерывная череда химических связей «консервирует» энергию, и работа системы не требует энергозатрат.

Общим для работы всех топоизомераз моментом является образование связи между одним краем каждого разрыва цепи ДНК («рваным» концом) и аминокислотным остатком тирозина в самом ферменте. Топоизомераза первого типа соединяет себя с одной (разорванной ею)-цепью, топоизомераза второго типа присоединяет к себе по одному концу каждой из двух разорванных ею цепей дуплекса ДНК. Топоизомеразы подразделяют на группы A и B по тому, с каким из фосфатов ДНК, 5′ или 3′, они образуют связь. Использование временной фосфодиэфирно-тирозиновой связи указывает на механизм действия фермента; он «перемыкает» фосфодиэфирную связь (связи) из ДНК на белок, манипулирует структурой одной или обеих цепей ДНК, а затем восстанавливает исходную связь (связи).

Все ферменты *E. coli* принадлежат к типу A (подсоединяются к 5′-фосфату). Это характерно для бактерий, у которых топоизомеразы типа B почти не встречаются. У эукариот, напротив, обнаружены топоизомеразы всех возможных типовых сочетаний (IA, IB, IIA и IIB).

Модель работы топоизомеразы IA проиллюстрирована на рис. 19.23. Фермент связывается с областью, где дуплекс ДНК разделен на отдельные цепи. Затем фермент рвет одну из цепей, пропускает сквозь разрыв другую цепь, и в завершение реакции разрыв заклеивает. Перенос связей с нуклеиновой кислоты на белок объясняет, каким образом ферменту удается функционировать без энергозатрат. Никакого необратимого гидролиза связей не происходит; их энергия «консервируется» благодаря реакциям переноса (оказывается запасенной в виде других связей). Анализ кристаллической структуры фермента подкрепляет данную модель.

В ходе топоизомеразной реакции порядок зацепления изменяется пошагово, размер шага равен единице. Всякий раз как одна цепь пропускается сквозь разрыв в другой, имеет место ΔL со значением +1. Рисунок иллюстрирует активность фермента с позиций наблюдения за движением каждой из цепей. В свободной сверхспирализованной молекуле такое изменение порядка зацеп-

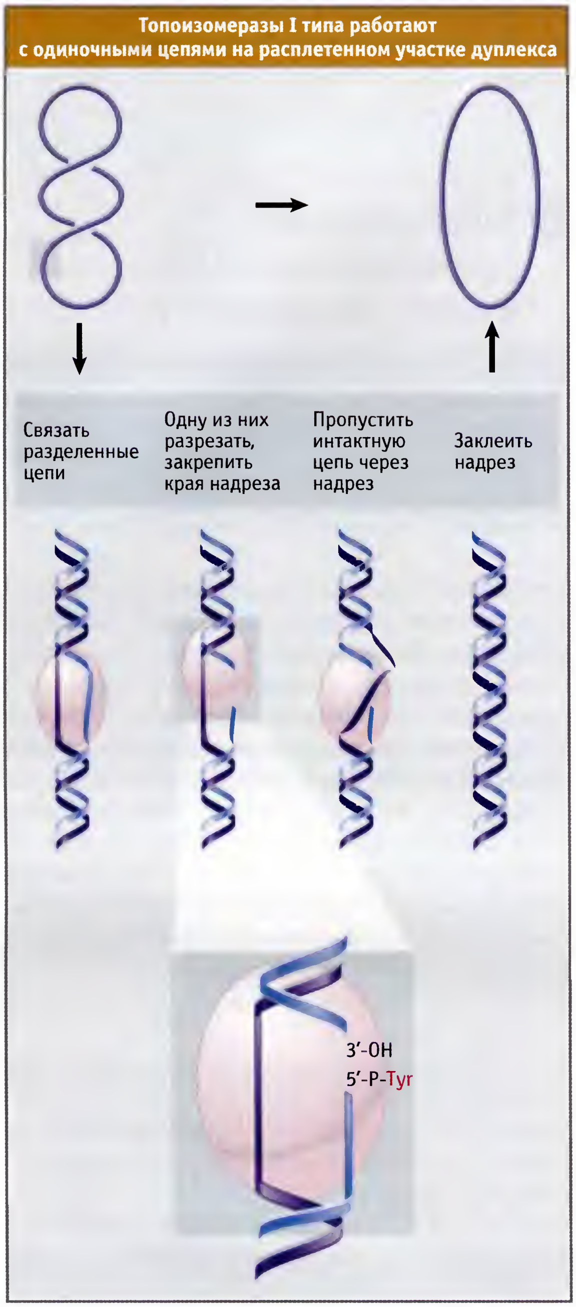


РИС. 19.23 У бактерий топоизомеразы I типа узнают частично расплетенные сегменты ДНК и пропускают одну цепь через разрез, сделанный в другой цепи

ления, несмотря на взаимозаменяемость W и T, следует принять как $\Delta W = +1$, то есть на один отрицательный сверхвиток меньше.

Реакция эквивалентна ротации, изображенной в нижней части рис. 19.21, с тем ограничением, что фермент переводит реакцию в режим серии шагов прохода одиночной цепи сквозь разрыв в другой цепи. (Для сравнения, бесконтрольное внесение надрезов в цепь сверхспирализованной молекулы позволяет этой цепи свободно вращаться, сбрасывая все напряжение множественными ротациями.)

Топоизомеразы первого типа умеют также **пропускать** один сегмент **одноцепочечной ДНК** сквозь другой. Эта реакция может приводить к образованию в ДНК **узлов** или катенировать две кольцевые молекулы друг за друга, соединяя их, как звенья цепочки. Непонятно, какое применение (и есть ли оно) у этих реакций *in vivo*.

Топоизомеразы второго типа в общем случае распускают и отрицательные, и положительные сверхвитки. Реакция требует расхода энергии, гидролиз одной молекулы АТФ обеспечивает один акт катализа. Как видно на РИС. 19.24, реакция опосредуется внесением двухцепочечного разрыва в один из дуплексов ДНК. Для этого производится разрез со смещением в 4 нуклеотида между цепями, и каждая субъединица димерного фермента прикрепляется к выступающему концу ДНК. Затем через разрыв пропускается другая область дуплекса. АТФ используется для последующего этапа лигирования/высво-

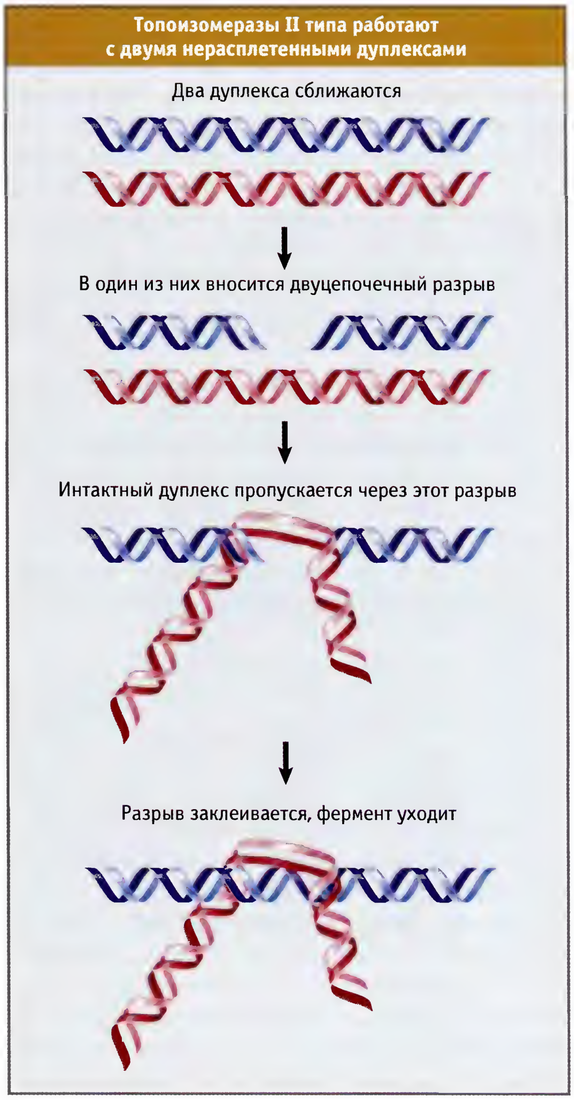


РИС. 19.24 Топоизомеразы II типа пропускают участок дуплекса через двухцепочечный разрыв в другом дуплексе

бождения, когда концы склеиваются обратно и дуплексы ДНК отпускаются. По этой причине ингибирование АТФазной активности фермента приводит к образованию «расщепляемого комплекса», содержащего разорванную ДНК.

Формальным следствием переноса двойной цепи является изменение порядка зацепления на величину, кратную двум. Также активность топоизомеразы второго типа может быть использована для образования или ликвидации катенированных кольцевых дуплексов и узлов ДНК.

Вероятно, эту реакцию можно представить как распознавание дуплексной ДНК, при котором фермент связывает любые два двухцепочечных сегмента, которые пересекаются. Гидролиз АТФ может использоваться для проведения фермента через конформационные преобразования, которые, в свою очередь, служат генератором мощности для проталкивания дуплекса ДНК сквозь разрыв в другом дуплексе. Согласно топологическим свойствам сверхспирализованной ДНК, взаимосвязь пересекающихся сегментов допускает удаление сверхвитков как из положительно, так и из отрицательно сверхспирализованных колец.

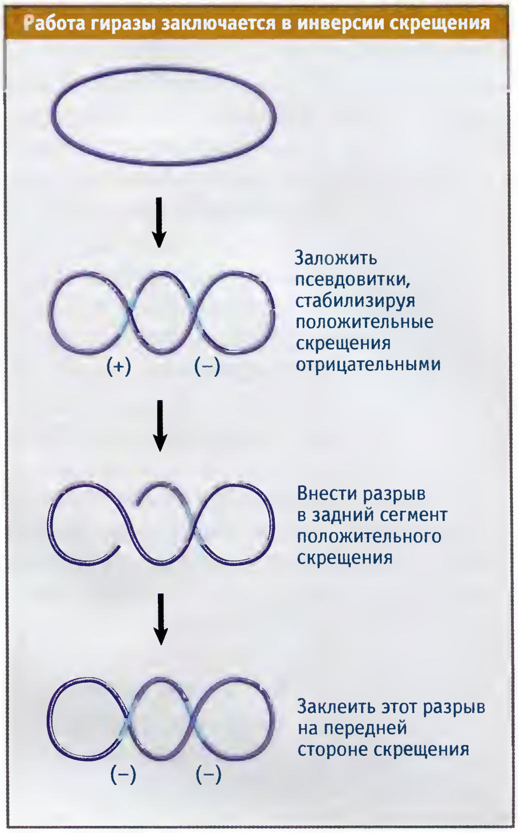


РИС. 19.25 ДНК-гираза может вносить в ДНК отрицательные сверхвитки посредством инверсии положительных скрещений сегментов ДНК

19.15 Гираза действует посредством инверсии витков

Ключевые понятия:

- Гираза *E.coli* — топоизомераза второго типа, использующая энергию гидролиза АТФ для внесения в ДНК отрицательных сверхвитков.

Бактериальная ДНК-гираза — это топоизомераза второго типа, способная вносить отрицательные сверхвитки в релаксированую закрытую кольцевую молекулу. ДНК-гираза связывается с кольцевым дуплексом и процессивно каталитически сверхспирализует его, методично внося сверхвитки в одну и ту же молекулу ДНК. Одна молекула ДНК-гиразы может за мину внести около 100 сверхвитков.

Сверхспирализованная форма ДНК обладает большей свободной энергией, чем релаксированная форма, сверхспирализация требует гидролиза АТФ. В отсутствие АТФ гираза может релаксировать отрицательные (но не положительные) сверхвитки, хоть и более чем в 10 раз медленнее скорости внесения сверхвитков.

ДНК-гираза *E. coli* — это тетрамер, состоящий из двух типов субъединиц, каждый из которых служит мишенью для антибиотиков, из которых чаще используемая налидиксовая кислота действует на GyrA, а новобиоцин — на GyrB. Эти лекарства ингибируют репликацию, что указывает на необходимость ДНК-гиразы для синтеза ДНК. Мутанты, резистентные к антибиотикам, помогли идентифицировать локусы, кодирующие субъединицы фермента.

При связывании с гиразой ДНК-субстрат обхватывает белковый тетрамер снаружи. Гираза закрывает примерно 140 пн ДНК от расщепления микрококковой нуклеазой. Модель **перемены знака** действия гиразы представлена на РИС. 19.25. Фермент связывает ДНК в пересеченной конфигурации, которая соответствует положительному сверхвитку. Это индуцирует образование компенсирующего отрицательного сверхвитка в несвязанной части молекулы. Затем фермент разрывает одну из двойных цепей на пересечении положительного сверхвитка, пропускает сквозь разрыв другой дуплекс, заклеивает разрыв.

Реакция эквивалентна инверсии знака сверхвитка: +1 оборот преобразуется в –1 оборот. Порядок зацепления, таким образом, изменился на $\Delta L = -2$, согласно правилу о кратности двум любого изменения значения порядка зацепления, связанного с пропусканием двойной цепи сквозь разрыв.

Затем гираза отпускает один из пересекающихся сегментов связываемого ею сверхвитка (теперь — отрицательного); это позволяет отрицательным оборотам перераспределиться вдоль ДНК (в виде изменения T или W, или обоих), и цикл начинается снова. Та же модель

топологических манипуляций лежит в основе образования конкатенанов и узлов.

По высвобождении инвертированного сверхвитка конформация гиразы меняется. Для проведения ферментом следующего цикла сверхскручивания, его исходная конформация должна быть восстановлена. Этот процесс называют **круговоротом фермента**. Полагают, что им движет гидролиз АТФ, поскольку замена АТФ негидролизуемым аналогом не позволяет гиразе вносить больше одной инверсии (в размере −2 сверхвитков) на субстрат. Таким образом, фермент нуждается в АТФ не для реакции сверхскручивания, но для перехода ко второму циклу. Новобиоцин препятствует течению АТФ-зависимых реакций гиразы, не давая АТФ связаться с субъединицей B.

АТФ-независимая реакция релаксации ингибируется налидиксовой кислотой. Это проясняет роль субъединицы A в реакции разрыва и воссоединения. Обработка гиразы налидиксовой кислотой приводит к «выходу» ДНК в виде фрагментов дуплекса, разрезанного поперек со смещением. Фрагменты имеют по краям свободную 3′-OH группу и 5′-одноцепочечный выступ в 4 нуклеотида, ковалентно сшитый с субъединицей A. Ковалентная сшивка хранит энергию фосфатной связи; эта энергия может быть использована для реакции лигирования — вот почему гираза может провести релаксацию в отсутствие АТФ. Сайты разрезов довольно специфичны, плотность их расположения составляет примерно один на 100 пн.

19.16 Специализированная рекомбинация задействует специфичные сайты

Ключевые понятия:

- Специализированная рекомбинация опосредована взаимодействием между определенными (необязательно гомологичными) сайтами.
- Фаг лямбда интегрирует в бактериальную хромосому посредством рекомбинации между сайтом фага и сайтом *att* на хромосоме *E.coli*.
- Фаг вырезается из хромосомы посредством рекомбинации между сайтами на концах линейного профага.
- Ген *int* фага лямбда кодирует интегразу, катализирующую реакцию интеграции.

В понятие специализированной (сайт-специфичной) рекомбинации входит реакция между двумя определенными сайтами. Длина сайтов-мишеней невелика (обычно от 14 до 50 пн). В одних случаях взаимодействующие сайты имеют одну и ту же последовательность, в других случаях они не гомологичны. Так, при интеграции свободной ДНК фага в хромосому бактерии или вырезании ее оттуда две рекомбинирующих последовательности отличаются. Та же реакция используется перед делением для регенерации мономерных кольцевых хромосом из димера, образованного в процессе общей (гомологичной) рекомбинации (см. разд. 17.7 «Для сегрегации хромосом может быть необходима сайт-специфичная рекомбинация»). В этом случае рекомбинирующие последовательности идентичны.

Ферменты, катализирующие сайт-специфичную рекомбинацию, обобщенно называют рекомбиназами, и их сейчас известно более сотни. Те из них, что участвуют в интеграции фага, а также родственные им ферменты относят к семейству интеграз. Члены семейства интеграз — прототип Int фага лямбда, Cre фага P1, дрожжевой фермент FLP (катализирующий инверсию хромосом).

Классическую модель **сайт-специфичной рекомбинации** можно проиллюстрировать на примере фага лямбда. Переходы ДНК лямбда между двумя ее жизненными формами заключают в себе два типа событий. Паттерн экспрессии генов регулируется согласно описанию в гл. 14 «Стратегии бактериофагов». Физическое состояние ДНК у умеренной (лизогенной) и литической форм бактериофага различается:

При литическом «образе жизни» ДНК лямбда существует в инфицированной бактерии как независимая кольцевая молекула.

В лизогенной форме фаговая ДНК составляет интегральную часть бактериальной хромосомы и называется **профагом**.

Переход между этими формами включает в себя сайт-специфичную рекомбинацию:

- Чтобы войти в лизогенное состояние, свободная ДНК лямбда должна быть вставлена в ДНК хозяина. Это явление называют **интеграцией**.
- Чтобы освободиться от умеренности (лизогении) и войти в литический цикл, ДНК профага должна быть выпущена из хромосомы. Это называют вырезанием, или **эксцизией**.

Интеграция и эксцизия достигаются рекомбинацией между бактериальной и фаговой ДНК по особым локусам, называемым сайтами прикрепления (***att***). Сайт прикрепления лямбда на бактериальной хромосоме в генетике бактерий обозначают как *att*$^{\lambda}$. Локус обнаруживают мутации, мешающие интеграции лямбда. В лизогенных штаммах локус *att*$^{\lambda}$ занят профагом **λ**. Если в хромосоме *E. coli* сайт *att*$^{\lambda}$ делетирован, то инфицирующий фаг лямбда все же может установить лизогению, интегрировав в другом месте, хотя эффективность такой реакции будет составлять <0,1% от частоты интеграции в *att*$^{\lambda}$. Эта неэффективная интеграция имеет место во **вторичных сайтах прикрепления**, последовательности которых напоминают истинную последовательность *att*.

При описании реакций интеграции/эксцизии бактериальный сайт прикрепления (*att*$^{\lambda}$) называют *attB*. По последовательности он состоит из компонентов *BOB′*. Сайт прикрепления на фаге, *attP*, состоит из компонентов *POP′*. На [illegible] дана схема реакции рекомбинации между этими сайтами. Последовательность O

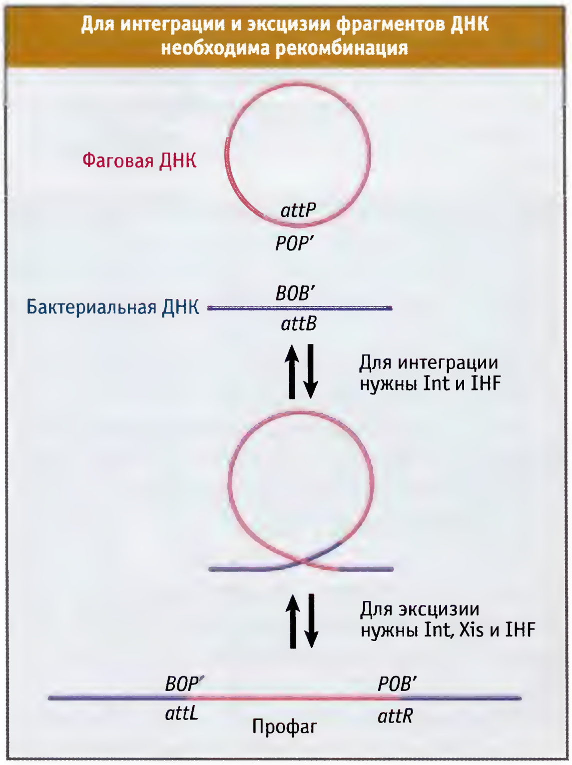


РИС. 19.26 После реципрокной рекомбинации между сайтами *attP* и *attB* кольцевая фаговая ДНК превращается в интегрированный профаг. Эксцизия профага обеспечивается реципрокной рекомбинацией между сайтами *attL* и *attR*

является общей для *attB* и *attP*. Ее называют **коровой** (центральной) **последовательностью**, и акт рекомбинации происходит внутри нее. Фланкирующие области *B*, *B′* и *P*, *P′* называют **плечами**; каждое со своей особой последовательностью. ДНК фага, замкнутую в кольцо, рекомбинация преобразует в линейную вставку на бактериальной хромосоме. Вставка содержит последовательность профага, ограниченную двумя новыми сайтами *att* — продуктами рекомбинации, — называемыми *attL* и *attR*.

Важным следствием такого устройства сайтов *att* является различие между парами последовательностей, участвующих в интеграции и эксцизии. Интеграция требует узнавания между *attP* и *attB*, в то время как эксцизия требует узнавания между *attL* и *attR*. Направленный характер сайт-специфичной рекомбинации контролируется оригинальностью каждого из рекомбинирующих сайтов.

Теоретически событие рекомбинации обратимо, но каждое из направлений реакции преобладает в разных условиях. Эта особенность важна для жизни фага, поскольку благодаря ей эксцизия и интеграция застрахованы от незамедлительного «поворота вспять».

Отличие пар сайтов, взаимодействующих при интеграции и эксцизии, отражается в разном наборе белков, опосредующих эти реакции:

- Для интеграции (*attB* × *attP*) требуется продукт фагового гена *int*, кодирующего фермент интегразу, а также бактериальный белок, называемый фактором интеграции хозяина (integration host factor, сокращенно IHF).
- Для эксцизии (*attL* × *attR*), кроме Int и IHF, требуется продукт фагового гена *xis*.

Таким образом, Int и IHF требуются для обеих реакций. Xis играет важную роль в контроле направления; он требуется для эксцизии, но тормозит интеграцию.

Похожая система, но с несколько упрощенными требованиями как к нуклеотидным последовательностям, так и к белковым компонентам, обнаружена у бактериофага P1. Кодируемая фагом рекомбиназа Cre катализирует рекомбинацию между двумя последовательностями-мишенями. В отличие от фага лямбда, у которого рекомбинирующие последовательности различны, у P1 они идентичны. Каждая состоит из 34 пн и называется *loxP*. Для реакции достаточно одной рекомбиназы; вспомогательные белки не требуются. Вследствие своей простоты и эффективности то, что теперь известно как «система Cre/*lox*», было адаптировано для использования в эукариотических клетках, для которых эта система стала одним из стандартных инструментов проведения сайт-специфичной рекомбинации.

19.17 Сайт-специфичная рекомбинация сопровождается разрывом и воссоединением

Ключевые понятия:

- Надрезы со смещением в 7 пн производятся в *attB* и *attP* и концы соединяются крест накрест.

Сайты *att* устроены в соответствии с различием требований к их последовательностям (*attP* гораздо длиннее *attB*). Функция *attP* требует протяженности в 240 пн, тогда как для функции *attB* достаточно 23 пн, заключенных между положениями −11 и +11, а с каждой стороны от коровой последовательности находится всего по 4 пн. Несоразмерность сайтов указывает на различие ролей *attP* и *attB* в рекомбинации, где *attP* несет дополнительную информацию, необходимую для различения его с *attB*.

Идет ли реакция по пути, при котором двойные цепи *attP* и *attB* перерезаются одновременно и затем обмениваются? Или обмен цепей производится по двухтактному механизму, в котором первый такт (обмен цепями) порождает, а второй такт (никование и лигирование) размыкает структуру Холлидея? Альтернативные варианты изображены на РИС. 19.27.

Найти ответ помогли эксперименты. Реакцию рекомбинации приостанавливали на промежуточных стадиях, используя «субстраты-самоубийцы», у которых коровая последовательность содержала ник

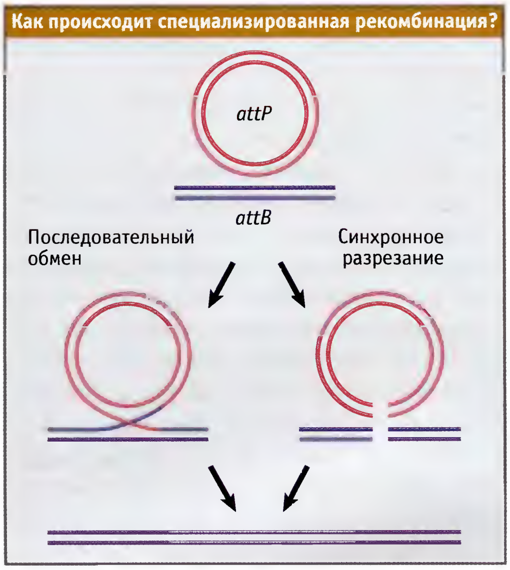


РИС. 19.27 По какой из двух схем идет специализированная (сайт-специфичная) рекомбинация между сайтами *attP* и *attB*?

Наличие такого ника препятствует процессу рекомбинации, что дает возможность идентифицировать молекулы, у которых рекомбинация началась, но не завершилась. Строение таких промежуточных продуктов позволяет считать, что обмен цепями происходит постепенно.

РИС. 19.28 иллюстрирует модель, где, если сайты *attP* и *attB* претерпевают одинаковые разрезы со смещением, комплементарные одноцепочечные концы подходят для перекрестной гибридизации. Расстояние между точками перехода цепей для лямбда равняется 7 пн, а при разрезе образуются 3′-фосфатные и 5′-гидроксильные концы. Для простоты на рисунке изображено, что перекрывающиеся «липкие» концы взаимно отжигаются (гибридизуются), но на самом деле весь процесс сродни событию рекомбинации с рис. 19.6. Соответствующие цепи каждого дуплекса разрезаются по одной и той же позиции, дуплексы обмениваются свободными 3′-концами, ветвь мигрирует на 7 пн вдоль участка гомологии, затем структура разделяется перерезанием другой пары цепей ДНК.

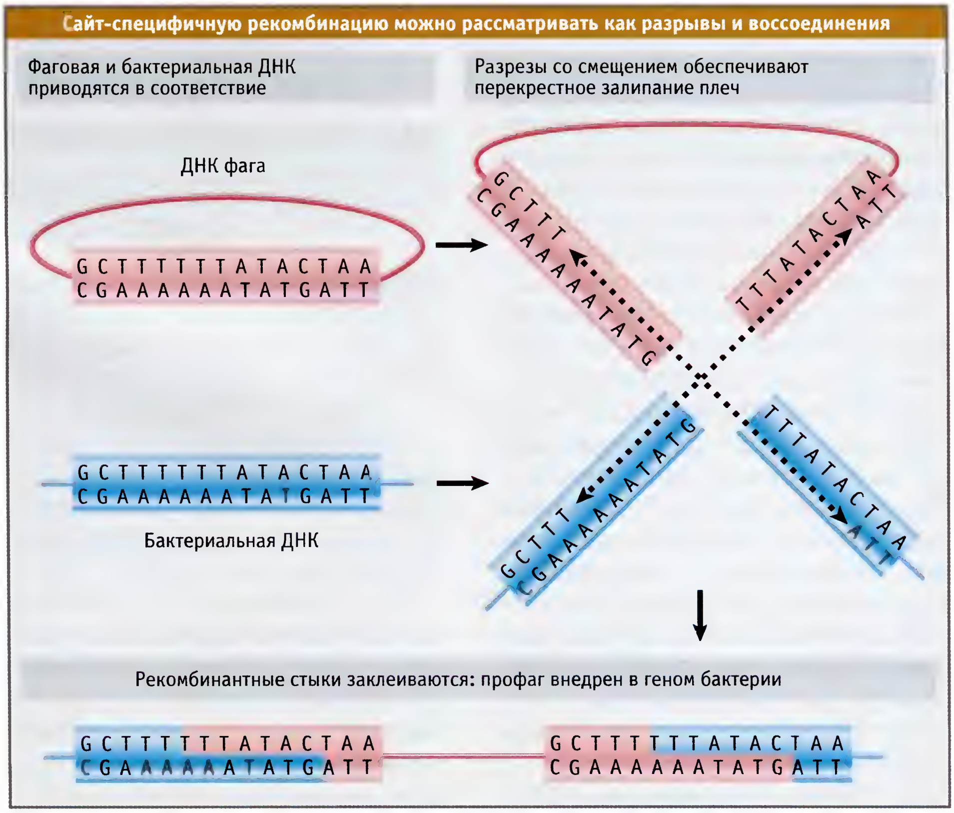


РИС. 19.28 Разрезы со смещением в идентичных участках элементов *attP* и *attB* создают возможность для перекрестного воссоединения концов с эффектом реципрокной рекомбинации

19.18 Сайт-специфичная рекомбинация напоминает топоизомеразную активность

Ключевые понятия:

- Интегразы родственны топоизомеразам, и реакция рекомбинации напоминает действие топоизомеразы, за исключением того, что «никованые» цепи *разных* дуплексов склеиваются вместе.
- Реакция «консервирует» энергию, используя каталитический тирозин в ферменте для разрушения фосфодиэфирной связи и привязки к «рваному» 3′-концу.
- Две субъединицы фермента связываются с каждым из сайтов рекомбинации и два таких димера сходятся, чтобы образовать комплекс, в котором происходит реакция переноса.

В своей работе интегразы, подобно топоизомеразам первого типа, используют механизм с внесением разрыва в одну цепь ДНК. (Различие состоит в том, что рекомбиназа затем соединяет «рваные» концы цепей крест-накрест, в то время как топоизомераза вносит разрыв, манипулирует концами, а затем соединяет концы по старым стыкам.) Механизм основан на присутствии четырех молекул рекомбиназы, каждая из которых разрезает одну из четырех цепей в двух рекомбинирующих дуплексах.

На РИС. 19.29 показана природа реакции, катализируемой интегразой. Фермент представляет сбой мономерный белок, обладающий активным сайтом, способным резать и лигировать ДНК. В ходе реакции тирозин атакует фосфодиэфирную связь. Из-за этого 3′-конец разрезанной цепи ДНК оказывается сшит фосфодиэфирной связью с тирозином фермента, а 5′-гидроксильный конец оказывается свободным.

С каждым сайтом рекомбинации связывается по две субъединицы фермента. Причем в каждом сайте лишь одна из субъединиц атакует ДНК. Симметрия системы гарантирует, что разорванными оказываются комплементарные цепи. Свободный 5′-гидроксильный конец каждого сайта атакует 3′-фосфотирозиновую сшивку другого сайта. В результате образуется структура Холлидея.

Разделение этой структуры происходит, когда две других субъединицы фермента (не участвовавших в первом цикле разрыва и воссоединения) воздействуют на другую пару комплементарных цепей.

Череда взаимодействий приводит к консервативному обмену цепями, при котором делеций или добавлений нуклеотидов в сайте обмена не происходит, и нет нужды в энергозатратах. Временная фосфотирозиновая сшивка между белком и ДНК хранит энергию разрушенной фосфодиэфирной связи.

На [illegible] можно видеть схему промежуточного продукта реакции, основанную на анализе его кристаллической структуры. («Отлов» промежуточного продукта оказался возможен благодаря использованию «суб-

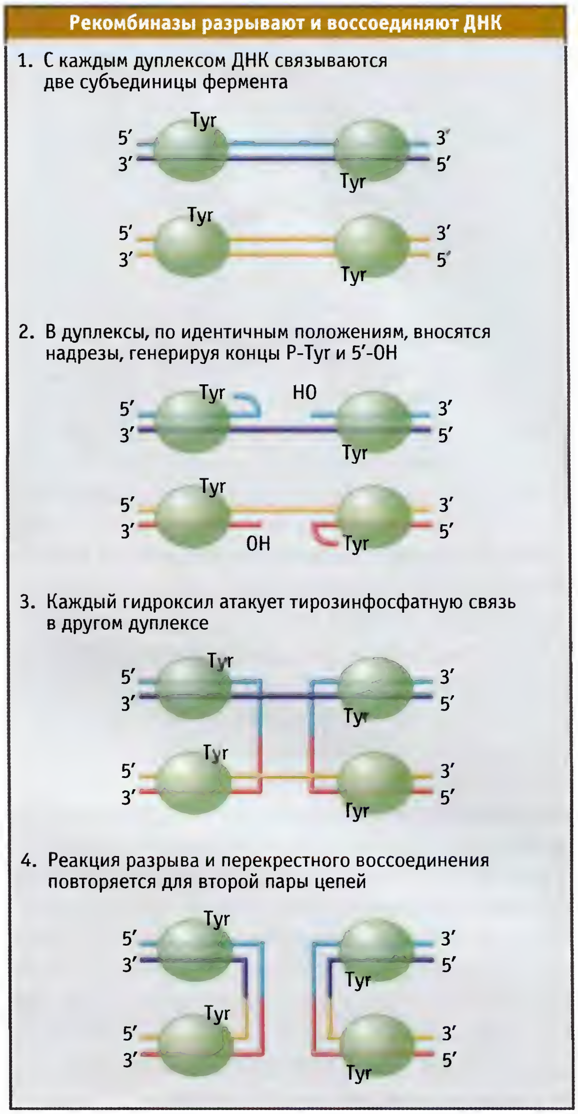


РИС. 19.29 Интегразы по механизму действия аналогичны топоизомеразам. В два дуплекса вносится по надрезу (с 3′-концевыми фосфатами, ковалентно пришитыми к тирозину фермента). Свободная гидроксильная группа каждой цепи атакует фосфотирозиновую связь гомологичной цепи другого дуплекса. Первый акт обмена цепями из двух, изображенных на рисунке, дает структуру Холлидея. Ее разделение происходит за счет повтора реакции для другой пары цепей

страта-самоубийцы», состоявшего из синтетического дуплекса ДНК с отсутствовавшей фосфодиэфирной связью, так что в результате атаки фермента свободный 5′-OH конец не образовывался.) В составе структуры Cre-*lox* комплекса видны две молекулы Cre, каждая из которых связана с отрезком ДНК в 15 пн. В центре симметрии молекула ДНК согнута под углом ~100°. Два таких комплекса сходятся в антипараллельной ориентации с образованием тетрамерной белковой структуры, связанной с двумя синапсированными молекулами ДНК. Обмен цепями имеет место в центральной полос-

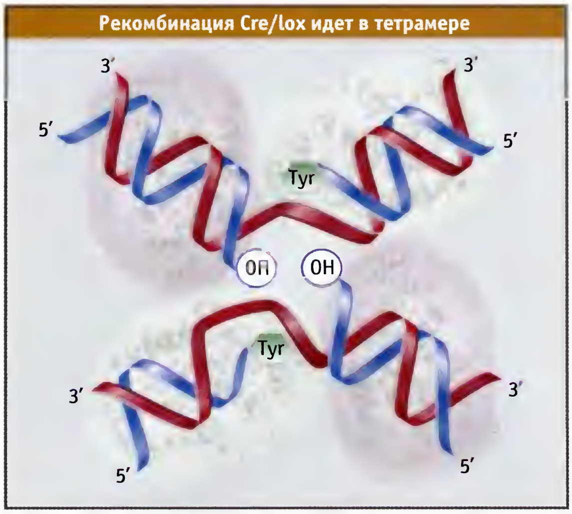


РИС. 19.30 В рабочем состоянии Cre-рекомбиназа представляет собой тетрамер, поскольку с каждым плечом сайта *loxA* связан один мономер фермента. Из четырех активных центров одновременно используется два; они действуют на комплементарные цепи двух сайтов ДНК

ти белковой структуры, которая вмещает шесть центральных нуклеотидов области перехода.

Тирозин, «ответственный» за разрезание ДНК по конкретному полусайту, несет в себе субъединица, прикрепленная к этому полусайту. Это явление называют *цис*-разрезом. Так же действуют интеграза Int и рекомбиназа XerD. В отличие от них рекомбиназа FLP режет ДНК в *транс*-ориентации. При таком механизме тирозин предоставляет *не* та субъединица, которая связана с этим полусайтом, а одна из остальных.

19.19 Рекомбинация фага лямбда происходит в интасоме

Ключевые понятия:

- Интеграция фага лямбда происходит в обширном комплексе, который, помимо прочих, содержит белок IHF хозяина.
- Реакция вырезания требует участия белков Int и Xis и узнает концы ДНК профага в качестве субстратов.

В отличие от рекомбинации в системе Cre/*lox*, где нужен лишь фермент и два сайта, рекомбинация фага лямбда происходит в обширной структуре и с различными компонентами для каждого из направлений реакции (интеграции либо эксцизии).

Белок хозяина, называемый IHF, требуется как для интеграции, так и для эксцизии. IHF весит 20 кДа и состоит из двух разных субъединиц, кодируемых генами *himA* и *himD*. Наличие IHF для клетки *E. coli* не очень существенно и не требуется при гомологичной бактериальной рекомбинации. Он является одним из нескольких белков, способных наматывать ДНК на свою поверхность. Препятствующие сайт-специфичной рекомбинации фага лямбда мутации в генах *him* могут быть супрессированы мутациями в *λint*, и это указывает на взаимодействие белков IHF и Int. Этими двумя белками можно провести сайт-специфичную рекомбинацию *in vitro*.

Реакция *in vitro* требует сверхспирализации *attP*, но не *attB*. При постановке реакции *in vitro* с участием двух сверхспирализованных молекул продукты реакции сохраняют почти ту же степень сверхспирализации. Это говорит об отсутствии свободных промежуточных продуктов, у которых могла бы иметь место ротация цепей, что с самого начала наводило на мысль о происхождении реакции через образование структуры Холлидея. Теперь нам известно, что реакция идет по механизму, типичному для ферментов этого класса и родственному механизму работы топоизомераз первого типа (см. разд. 19.18 «Сайт-специфичная рекомбинация напоминает топоизомеразную активность»).

Белок Int может связываться с ДНК двумя разными способами. Его C-концевой домен ведет себя подобно рекомбиназе Cre. Он связывает инвертированные сайты в составе коровой последовательности, располагаясь для проведения реакций разрезания и лигирования по положениям, обозначенным на РИС. 19.31. N-концевой домен связывает сайты с разной консенсусной последовательностью, расположенные в плечах *attP*. Это связывание отвечает за агрегацию отдельных элементов в интасому. Вероятно, два домена Int связывают ДНК одновременно, таким способом подтягивая плечи к центральной последовательности.

Белок IHF связывает в *attP* последовательности размером ~20 пн. Сайты связывания IHF примыкают к сайтам

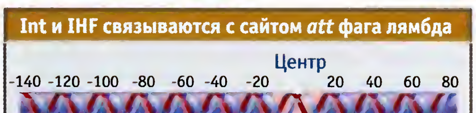


РИС. 19.31 Int и IHF связывают разные сайты элемента *attP*. Сайты связывания Int в центральной области элемента содержат точки разрезов

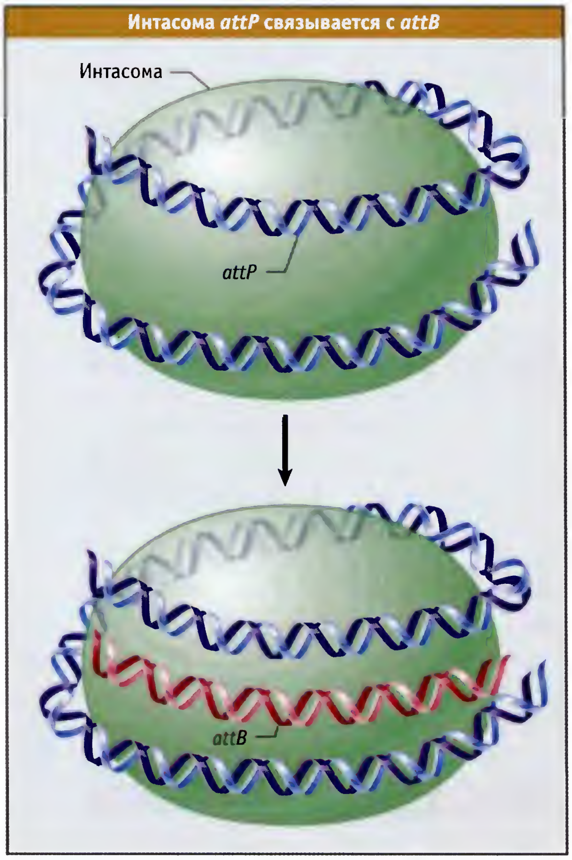


РИС. 19.32 **Множество копий белка Int может придавать элементу *attP* вид интасомы, инициирующей сайт-специфичную рекомбинацию при узнавании в свободной ДНК сайта *attB***

связывания Int. Препятствующий интеграции белок Xis связывает *attP* по двум близким сайтам, защищая участок в почти 40 пн ДНК. Все вместе Int, Xis и IHF покрывают практически весь *attP*. Связывание с Xis изменяет организацию ДНК, делая ее, как субстрат, инертной для интеграции.

Когда Int и IHF связываются с *attP*, они порождают комплекс, в котором все сайты связывания стянуты вместе на поверхности белка. Для образования такой **интасомы** необходима сверхспирализация *attP*. У *attB* для связывания есть лишь два сайта Int в коровой последовательности. При этом Int не связывает *attB* в виде свободной ДНК. Посредником при «захвате» *attB* служит интасома. Это схематично обозначено на РИС. 19.32.

В соответствии с этой моделью, исходное узнавание между *attP* и *attB* не зависит напрямую от гомологии ДНК, а определяется способностью белков Int узнавать обе последовательности *att*. Затем происходит сближение двух сайтов *att* в ориентации, предопределенной строением интасомы. С этого момента гомология последовательностей уже важна, поскольку требуется для обмена цепями.

Асимметрию реакций интеграции и эксцизии показывает тот факт, что Int может образовать похожий комплекс с *attR* только после добавления Xis. Этот комплекс может образовать пару с конденсированным комплексом, который Int образует в *attL*. Для этой реакции белок IHF не нужен. Принципиальное разрличие интеграции/эксцизии лямбда и реакций рекомбинации, катализируемых Cre или Flp, состоит в том, что в Int-катализируемых реакциях связывание регуляторных последовательностей в плечах сайтов-мишеней приводит к сгибанию ДНК, обусловливающему взаимодействие между плечами и коровой (центральной) последовательностью и доводящие реакцию до финала. Поэтому каждая из реакций лямбда необратима в отличие от обратимой рекомбинации под контролем Cre или Flp. Анализ кристаллической структуры тетрамеров λ-Int показывает, что, как и в других рекомбиназах, у тетрамера есть две активных и две неактивных субъединицы, которые по ходу рекомбинации меняются ролями. Аллостерические взаимодействия внутри белка, переключаемые эффектами связывания с плечами ДНК, контролируют смену структурных состояний тетрамера и, тем самым, течение реакции.

Сложность сайт-специфичной рекомбинации во многом вызвана потребностью регулировать реакцию таким образом, чтобы интеграции оказывалось предпочтение при вхождении вируса в лизогенное состояние, а эксцизии — при вхождении профага в литический цикл. Соответствие направления реакции жизненного цикла достигается контролем концентрации Int и Xis.

19.20 Молчащие и активные локусы переключаются при смене типа спаривания у дрожжей

Ключевые понятия:

- **Локус MAT, отвечающий за тип спаривания у дрожжей, может иметь генотип *MAT*a или *MAT*α.**
- **Дрожжи с доминантным аллелем *HO* переключают тип спаривания.**
- **Аллель, находящийся в локусе *MAT*, называют активной кассетой.**
- **Есть еще две молчащих кассеты, *HML*α и *HMR*a.**
- **Переключение происходит, если *MAT*a заменяется *HML*α или если *MAT*α заменяется *HMR*a.**

Дрожжи *S. cerevisiae* могут распространяться как в гаплоидном, так и в диплоидном состоянии. Переход между состояниями происходит при спаривании (слияние гаплоидных клеток с образованием диплоидной) и при споруляции (мейоз диплоидных клеток с образованием гаплоидных спор). Способность вступать в один из этих процессов определяется типом спаривания штамма, который бывает **a** или α. Гаплоидные клетки типа **a** могут спариваться лишь с гаплоидными клетками типа α с образованием

диплоидных клеток типа **a**/α. Эти диплоидные клетки могут спорулировать с восстановлением гаплоидных спор обоих типов.

Поведение при спаривании определяется генетической информацией, находящейся в локусе *MAT*. Клетки, которые по этому локусу несут аллель *MAT***a**, относят к типу **a**; клетки, несущие аллель *MAT*α, относят к типу α. Клетки противоположных типов спаривания узнают друг друга благодаря секреции феромонов: клетки α секретируют небольшой полипептид — α-фактор; клетки **a** секретируют **a**-фактор. Клетка одного типа спаривания несет на своей поверхности рецептор к феромону противоположного типа. Когда клетки **a** и α встречают друг друга, феромоны действуют на рецепторы, что приводит к обоюдному аресту клеточного цикла в фазе G1 и разнообразным морфологическим изменениям. При успешном спаривании за арестом клеточного цикла следует слияние клеток и их ядер с образованием диплоидной клетки **a**/α.

Спаривание — симметричный процесс, инициируемый взаимодействием феромона, секретированного клеткой одного типа, с рецептором на поверхности клетки другого типа. Гены, участие которых в клеточном пути ответа при конкретном типе спаривания является уникальным, — это гены рецепторов. Взаимодействие фактора **a** с рецептором запускает в клетке α тот же путь ответа, что и взаимодействие фактора α с рецептором в клетке **a**. Мутации, отключающие разные этапы этого пути (общего для двух клеточных типов), одинаково проявляются у **a** и α. Путь представляет собой проведение сигнала каскадом белковых взаимодействий и приводит к синтезу продуктов, вносящих необходимые изменения в клеточную морфологию и экспрессию генов для того, чтобы произошло спаривание.

Информация о сигнальном пути спаривания дрожжевых клеток была во многом почерпнута из свойств мутантных **a** и/или α клеток, неспособных к спариванию. Гены, идентифицируемые такими мутациями, называют *STE* (от слова sterile, стерильный). Мутации в генах феромонов и рецепторов специфичны для **a** или α, в то время как мутации в других генах *STE* исключают спаривание у клеток обоих типов. Это объясняется тем, что в клетках **a** и α за взаимодействием фактора с рецептором следуют одинаковые события.

Некоторые штаммы дрожжей обладают замечательной способностью к переключению типа спаривания. Эти штаммы-носители доминантного аллеля *HO* часто меняют тип своего спаривания — по разу в каждом поколении. Штаммы с рецессивным аллелем *ho* обладают стабильным типом спаривания, который подвержен изменению с частотой лишь ~10^{-6}.

Присутствие *HO* заставляет генотип популяции дрожжей меняться. Безотносительно исходного типа спаривания, в течение буквально нескольких поколений оба типа спаривания представлены огромным числом клеток, что приводит к образованию диплоидов *MAT***a**/*MAT*α, берущих верх над популяцией. Продукция гаплоидной популяцией стабильных диплоидов может рассматриваться как *raison d'être* для переключения.

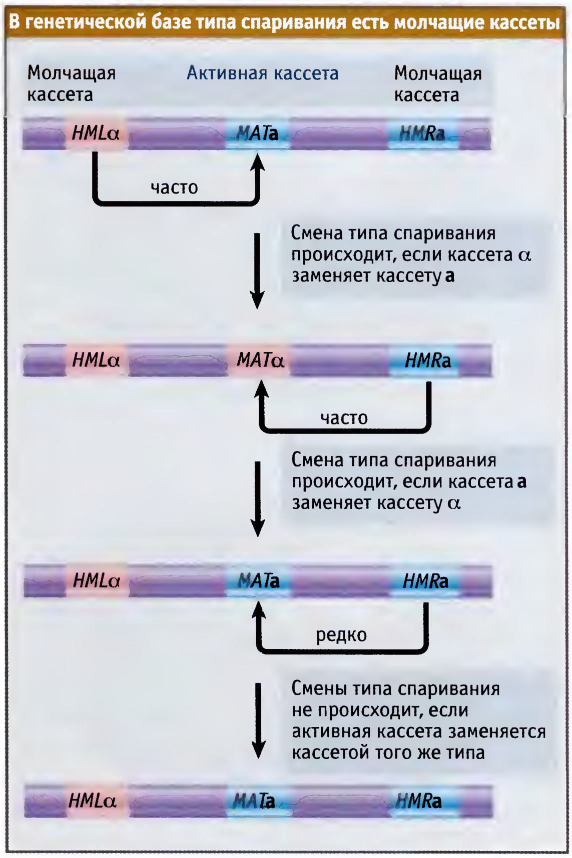


РИС. 19.33 Смена типа спаривания происходит, когда молчащая кассета заменяет собой активную кассету противоположного генотипа. Соответственно, перестановка a и a, или α и α, не приводит к смене типа спаривания

Само существование переключения указывает на то, что все клетки содержат потенциальную информацию как для *MAT***a**, так и для *MAT*α, но экспрессируют лишь один тип. Откуда поступает информация о смене типа спаривания? Для переключения нужны два дополнительных локуса: *HML*α для переключения на *MAT*α, *HMR***a** для переключения на *MAT***a**. Эти локусы лежат на той же хромосоме, что и *MAT*. Локус *HML* находится слева от *MAT*, *HMR* — справа (оба — довольно далеко от *MAT*).

Кассетная модель спаривания проиллюстрирована на РИС. 19.33. Она основана на положении, что в *MAT* есть *активная кассета* — либо α, либо **a**. В *HML* и *HMR* есть *молчащие кассеты*. В общем случае *HML* несет кассету α, а *HMR* — кассету **a**. Все кассеты несут в себе информацию о типе спаривания, но только активная кассета в *MAT* подлежит экспрессии. Переключение типа спаривания происходит, когда активная кассета заменяется ранее молчавшей, и эта новоинсталлированная кассета начинает экспрессироваться.

Переключение нереципрокно; аллель в *MAT* замещается копией из *HML* или *HMR*. Об этом говорит судьба мутаций по локусу *MAT*. При переключении типа спаривания они безвозвратно теряются — копия из MAT не вступает в перестановку с замещающим ее материалом.

Если мутирует молчащая копия гена в *HML* или *HMR*, то при переключении мутантный аллель вставляется в локус *MAT*. Остающаяся при этом в *HML* или *HMR* мутантная копия может пройти через неограниченное число переключений. Как при репликативной транспозиции, элемент-донор порождает новую копию на сайте реципиента, оставаясь при этом нетронутым.

Переключение типа спаривания на уровне ДНК является направленным событием, при котором на одного потенциального реципиента (*MAT*) приходится два потенциальных донора (*HML* и *HMR*). Под переключением обычно подразумевается замена *MAT***a** копией *HML*α, или замена *MAT*α копией *HMR***a**. С частотой от 0,8 до 0,9 переключение меняет тип аллеля в *MAT* на противоположный. Это определяется фенотипом клетки. Клетки с фенотипом **a** отдают предпочтение *HML* в качестве донора; клетки с фенотипом α отдают предпочтение *HMR*.

В переключении типа спаривания задействовано несколько групп генов. В дополнение к генам, напрямую детерминирующим тип спаривания, сюда относятся гены, необходимые для репрессии молчащих кассет переключения типа спаривания, исполнения ряда других функций.



РИС. 19.34 **Молчащие кассеты обладают той же нуклеотидной последовательностью, что и активные кассеты того же типа, за исключением делеции краев фланкирующих последовательностей у HMRa. Кассеты разных типов спаривания имеют похожие последовательности: a и α отличаются только областью Y**

Сравнив последовательности двух молчащих кассет (*HML*α и *HMR***a**) с последовательностями двух типов активной кассеты (*MAT***a** и *MAT*α), мы можем вычленить последовательности, определяющие тип спаривания. Организацию локусов резюмирует РИС. 19.34. Каждая кассета содержит общие последовательности, фланкирующие центральную область, разную для **a** и α. Эту область называют соответственно *Y***a** или *Y*α. Последовательности, фланкирующие область *Y*, практически идентичны, но в *HMR* они короче. Активная кассета в *MAT* транскрибируется с промотора внутри области *Y*.

19.21 Локус *MAT* кодирует белки-регуляторы

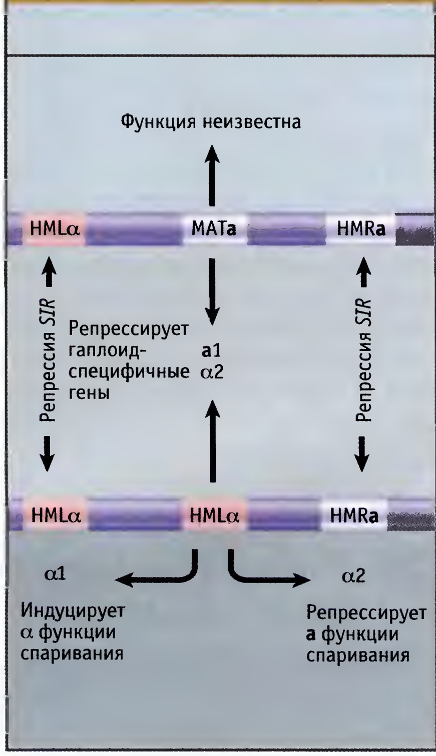
Ключевые понятия:

- **В гаплоидных клетках типа α, *MAT*α активирует гены, определяющие α-специфичные функции, необходимые для спаривания, и инактивирует гены, обусловливающие тип спаривания a.**
- **В клетках типа a *MAT*a не требуется.**
- **В диплоидных клетках продукты генов a*1* и α*2* кооперативно репрессируют гаплоид-специфичные гены.**

Основное назначение локуса *MAT* — контроль экспрессии генов феромонов и рецепторов, а также ряд других причастных к спариванию функций. *MAT*α кодирует два белка, α1 и α2. *MAT***a** кодирует единственный белок, **a***1*. Белки **a** и α напрямую контролируют экспрессию множества генов-мишеней, осуществляя как положительную (позитивную, в сторону усиления экспрессии), так и отрицательную (негативную, в сторону ослабления экспрессии) регуляцию. В гаплоидных клетках они действуют поодиночке, в диплоидных — сопряженно. Их взаимодействия по отношению к трем группам генов-мишеней резюмированы в правой части таблице на РИС. 19.35:

- Гены с **a**-специфичными функциями конститутивно (нерегулируемо) экспрессируются в клетках **a**. В клетках α они репрессированы (их экспрессия подавлена). К **a**-специфичным генам относятся структурный ген фактора **a** и *STE2*, кодирующий рецептор к фактору α. Таким образом, фенотип **a** ассоциирован с готовностью к узнаванию феромона, выделяемого клеткой с противоположным типом спаривания.
- Гены с α-специфичными функциями индуцированы в клетках α, но не экспрессируются в клетках **a**. К ним относятся структурный ген фактора α и ген рецептора к фактору **a**, *STE3*. Снова экспрессия феромона одного типа ассоциирована с экспрессией рецептора к феромону противоположного типа.

Транскрипция многих генов зависит от типа спаривания



	гены a	гены α	гаплоидные функции
a гаплоид	конститутивная экспрессия	нет экспрессии	конститутивная экспрессия
диплоид	нет экспрессии	нет экспрессии	репрессия
α гаплоид	репрессия	индукция	конститутивная экспрессия

РИС. 19.35 **В диплоидных клетках белки a1 и α2 совместно подавляют функции, характерные только для гаплоидных клеток. В гаплоидных клетках a идет конститутивная экспрессия функций спаривания a. В гаплоидных клетках α белок α2 репрессирует функции спаривания a, а белок α1 индуцирует экспрессию функций спаривания α**

- К генам с гаплоид-специфичными функциями относятся те, которые нужны для транскрипции феромонов и рецепторов, а также причастный к переключению ген *HO* и репрессор споруляции *RME*. Они конститутивно экспрессируются в гаплоидных клетках обоих типов, но репрессированы в диплоидных клетках **a**/α. Гены с **a**- и α-специфичными функциями тоже не экспрессируются в диплоидных клетках.

Теперь можно рассмотреть функции регуляторов и их мишеней с точки зрения функций *MAT* в гаплоидных и диплоидных дрожжевых клетках, как показано в левой части рис. 19.35. Типы спаривания **a** и α регулируются по разным механизмам:

- В гаплоидных клетках **a** гены, обслуживающие спаривание, экспрессируются конститутивно (по умолчанию). Функции продуктов *MAT***a** в клетке **a** (если они есть) неизвестны. Возможно, *MAT***a** требуется лишь для репрессии гаплоидных функций в диплоидной клетке.
- В гаплоидных клетках α белок α1 запускает экспрессию α-специфичных генов, чьи продукты нужны для спаривания по типу α. Другой продукт *MAT*α, белок α2, подавляет экспрессию генов, ответственных за спаривание по типу **a**, связываясь с последовательностью оператора, предлежащей генам-мишеням.
- В диплоидных клетках белки **a***1* и α2 кооперативно репрессируют гаплоид-специфичные гены. Они объединяются, чтобы распознать последовательность оператора, отличную от последовательности-мишени α2.

Способность белков α2, **a***1* и α1 регулировать транскрипцию опирается на ряд интересных белок-белковых взаимодействий. Паттерны контроля генов в клетках **a**, в клетках α и в диплоидных клетках резюмирует РИС. 19.36.

Во многих из этих взаимодействий участвует белок фактор транскрипции феромонового рецептора (PTRF). Он связывается с пи-боксом (P box) — короткой консенсусной последовательностью. Роль PTRF в регуляции генов может быть весьма разносторонней, поскольку пи-боксы встречаются по множеству положений. Некоторые пи-боксы нужны для активации генов; в других локусах PTRF нужен как репрессор. Характер влияния, судя по всему, может зависеть от других белков, которые связывают ДНК по соседству от пи-бокса.

Специфичную для гаплоидных клеток **a** активацию генов PTRF обеспечивает самостоятельно. В клетке **a** его одного достаточно для их экспрессии.

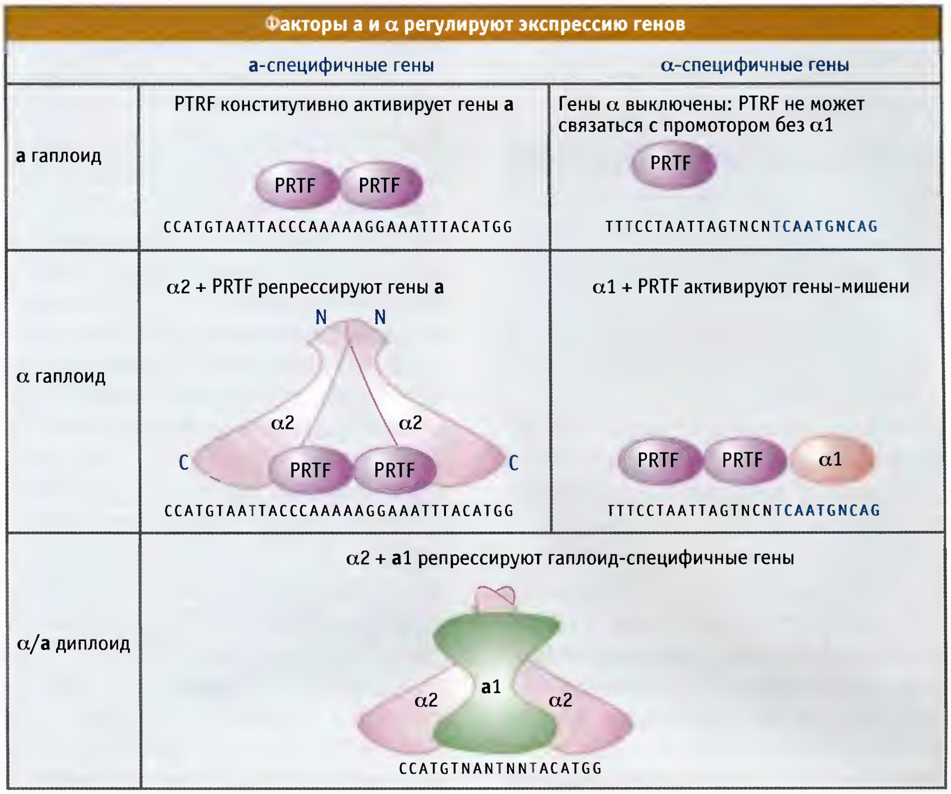


РИС. 19.36 **Различные комбинации белков PTRF, a1, α1 и α2 активируют либо репрессируют конкретные группы генов в соответствии с типом спаривания, присущим клетке**

В гаплоидной клетке α **a**-специфичные гены репрессируются сопряженной активностью PTRF и белка α2. У белка α2 есть два домена. С-концевой домен связывается с короткими взаимно-инвертированными элементами по краям консенсусной последовательности оператора длиной в 32 пн. Это связывание, однако, не вызывает репрессии. N-концевой домен, необходимый для репрессии, отвечает за контакты с PTRF. Сайтом связывания PTRF служит пи-бокс в центре последовательности оператора. На деле, α2 с PTRF кооперативно связывают ДНК.

Для экспрессия α-специфичных генов требуется другой небольшой белок, активатор α1, размером в 175 аминокислот. α-специфичную транскрипцию обусловливают, действуя в *цис*-положении, последовательности длиной в 26 пн. Эти 26 пн можно подразделить на две части. Первые 16 пн образуют пи-бокс, с которым связывается PTRF, смежные 10 пн образуют сайт связывания для α1. Фактор α1 связывается с ДНК только при наличии PTRF. Ни один из этой пары белков не может самостоятельно связаться с боксом-мишенью, вместе же они могут связать ДНК, предположительно, при поддержке белок-белковых взаимодействий.

В гаплоидных клетках **a** α-специфичные гены выключены по умолчанию, поскольку PTRF, в отсутствие белка α1, не может самостоятельно связаться с промотором, чтобы их активировать.

Белок α2 может также кооперировать с белком **a**1. Их комбинация опознает другой оператор. У этого оператора внешние взаимно-инвертированные участки предназначены для узнавания С-концевым доменом α2, но между ними лежит не пи-бокс, а другая, более короткая, последовательность. В диплоидных клетках комбинация **a**1/α2 репрессирует гены с таким мотивом.

Основной вывод состоит в том, что фенотип каждого клеточного типа (гаплоидного **a**, гаплоидного α или диплоидного **a**/α) определяется комбинацией белков **a** и α. Одним аспектом является различение гаплоидного и диплоидного фенотипов; другим — различение между гаплоидными фенотипами **a** и α. Последнее распространяется на экспрессию генов, соответствующих конкретному типу спаривания, и на определение направления смены типа спаривания (см. рис. 19.33). Клетки *MAT***a** активируют у себя энхансер рекомбинации на левом плече хромосомы III, который усиливает рекомбинацию на участке в 40 тпн, содержащем *HML*. В клетках *MAT*α левый край хромосомы III инактивирован.

19.22 Молчащие кассеты в *HML* и *HMR* репрессированы

Ключевые понятия:

- *HML* и *HMR* репрессированы с участием элементов-сайленсеров.
- В число локусов, необходимых для поддержания репрессии *HML* и *HMR*, входят *SIR1-4*, *RAP1* и гены гистона Н4.
- Связывание комплекса ORC с сайленсерами необходимо для инактивации.

При изучении карты транскрипции (см. рис. 19.34) обнаруживается интересная вещь: и для *MAT***a**, и для *MAT*α транскрипция инициируется в пределах области *Y*. При этом экспрессируется всегда только локус *MAT*, хотя такие же области *Y***a** и *Y*α есть и в молчащих кассетах локусов *HMR* и *HML* соответственно. Это означает, что регуляция экспрессии не ограничивается компетенцией сайта связывания белковых факторов в зоне промотора. *Функциональное отличие MAT от HML и HMR должно быть опосредовано сайтом, лежащим за пределами кассет.*

Делеционный анализ показал, что для подавления экспрессии как в *HML*, так и в *HMR* нужны сайты по обе стороны каждого из этих локусов. Эти сайты называют **сайленсерами** (silencers). Сайты по левую сторону кассеты называют E-сайленсерами, сайты по правую сторону называют I-сайленсерами. Эти сайты могут осуществлять контроль на расстоянии до 2,5 тпн от промотора и в любой ориентации. Они ведут себя как негативные энхансеры (энхансеры — это удаленные от промотора элементы, активирующие транскрипцию; см. разд. 24.15 «Энхансеры несут двунаправленные элементы, которые способствуют инициации»).

Каким должен быть подход к поиску генов, ответственных за состояние молчания кассет? Следовало бы ожидать, что продукты этих генов действуют на сайленсеры. Удобство поиска мутаций в таких генах обусловлено тем обстоятельством, что, если мутация позволяет экспрессироваться молчащей в норме кассете в *HMR* и *HML*, это приводит к сочетанию функций **a** и α в гаплоидной клетке, ведущей себя при этом как диплоидная клетка *MAT***a**/*MAT*α.

Было идентифицировано несколько локусов, мутации по которым отменяют сайленсинг, приводя к экспрессии *HML* и *HMR*. Сначала были открыты четыре локуса *SIR* (сокращение от silent information regulators, «регуляторы молчащей информации»). Все четыре *SIR* дикого типа необходимы для поддержания *HML* и *HMR* в репрессированном состоянии; мутация по любому из четырех локусов с образованием аллеля *sir*⁻ имеет два эффекта. Во-первых, *HML* и *HMR* оба могут транскрибироваться; во-вторых, обе молчащих кассеты становятся мишенями для замещения при переключении типа спаривания. Таким образом, одна и та же регуляторная схема участвует в репрессии молчащей кассеты и одновременно защищает ее от исполнения роли реципиента, замещаемого другой кассетой.

Также в сайленсинге участвуют продукты локуса *RAP1* (который, помимо этого, требуется для поддержания теломерного гетерохроматина в инертном состоянии) и гены гистона Н4. Делеции N-конца или отдельные точечные мутации гистона Н4 активируют молчащие кассеты. Последствия этих мутаций могут быть преодолены либо введением новых мутаций в *SIR3*, либо повышенной экспрессией *SIR1*. Это говорит о возможности особого взаимодействия гистона Н4 с белками SIR.

Согласно модели, основанной на этих данных, белки SIR препятствуют экспрессии генов, в частности *HML* и *HMR*, воздействуя на структуру хроматина. Мутации в белках SIR точно так же действуют на гены, которые были инактивированы близостью теломерного гетерохроматина. Вероятно, в общем случае, белки SIR, взаимодействуя с гистонами, образуют гетерохроматиновые (инертные) структуры (см. разд. 31.3 «Существование гетерохроматина обусловлено взаимодействиями гистонов с другими белками»).

Существует интересная взаимосвязь между репрессией в сайленсерах и репликацией ДНК. Каждый сайленсер содержит последовательность *ars* (ориджин репликации). С *ars* связывается ORC (сокращение от origin recognition complex, комплекс узнавания ориджина), причастный к инициации репликации. Мутации в генах *ORC* препятствуют сайленсингу, показывая, что связывание белка ORC сайленсером необходимо для сайленсинга.

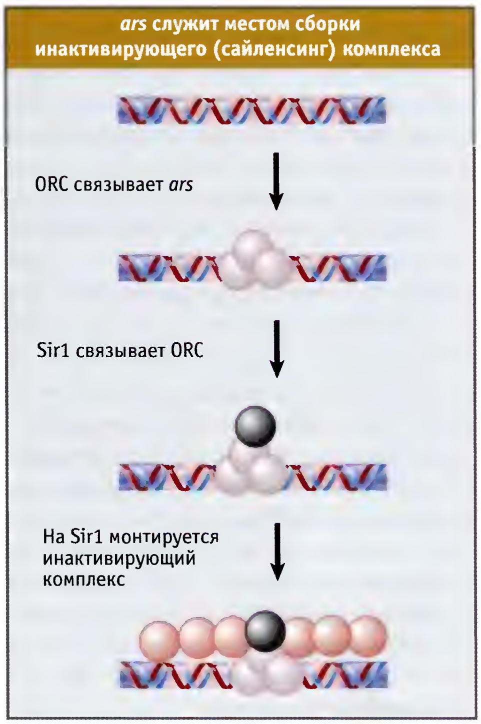


Рис. ... Для сайленсинга в HMR нужен Sir1

Существует два отдельных типа связи между сайленсингом и аппаратом репликации:

- необходимость присутствия Sir1, и
- участие репликации.

Если локализовать белок Sir1 в сайленсере (сшив его с другим белком, связывающим там ДНК), то связывание ORC для репрессии не понадобится. Это означает, что роль ORC в этом проекте — только доставка Sir1, инициация репликации от него не зависит. Роль ORC иллюстрирует РИС. 19.37. Вероятно, он обеспечивает всему процессу сайленсинга «инициирующий центр»: предоставляет структурную основу для посадки Sir1, который затем рекрутирует остальные белки SIR. Таким образом, роль ORC здесь совсем другая, чем в репликации.

Чтобы сайленсинг установился, необходимо прохождение S-фазы клеточного цикла. Это не подразумевает инициации репликации на *ars* в сайленсере. Эффект может определяться прохождением репликативной вилки через сайленсер, изменяющим местную структуру хроматина.

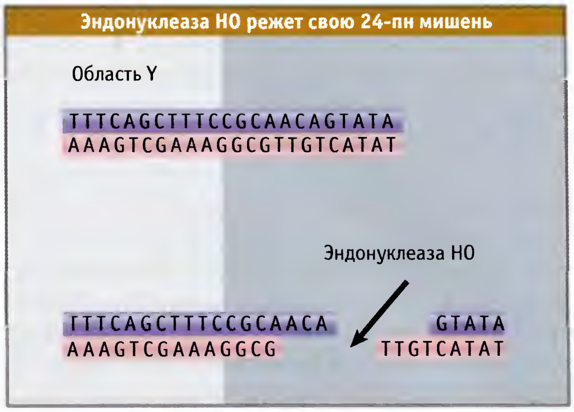


РИС. 19.38 Эндонуклеаза HO режет локус *MAT* справа от области Y. Получаются липкие концы с выступом в четыре основания

19.23 Однонаправленный перенос инициируется локусом-реципиентом (*MAT*)

Ключевые понятия:

- Переключение типа спаривания инициируется двухцепочечным разрывом, который эндонуклеаза HO вносит в локус *MAT*

Тип спаривания переключается за счет конверсии гена, в ходе которой сайт-реципиент (*MAT*) обретает последовательность донора (*HML* или *HMR*). Сайты этого локуса, важные для транспозиции, были идентифицированы по мутациям *MAT*, препятствующим переключению. Однонаправленная природа процесса засвидетельствована отсутствием подобных мутаций по *HML* или *HMR*.

Сайт по правой границе *Y* в локусе *MAT* «отмечен» мутациями как принципиально важный для переключения. Его границы демонстрирует анализ расположения этих точечных мутаций относительно сайта переключения (это можно установить, исследуя результаты редких событий переключения, несмотря на мутацию). Некоторые мутации лежат внутри замещаемой области (и, следовательно, исчезают из *MAT* после переключения), в то время как другие находятся чуть в стороне от замещаемой области (продолжая и в дальнейшем мешать переключению). Таким образом, для акта переключения нужны последовательности, лежащие как вне, так и внутри замещаемой области.

Переключение инициируется двухцепочечным разрывом вблизи границы *Y-Z*, по сайту, чувствительному к действию ДНКазы (это общая черта сайтов, задействованных в инициации транскрипции или рекомбинации). Сайт узнает эндонуклеаза, кодируемая локусом *HO*. Справа и в непосредственной близости от границы *Y* эндонуклеаза HO вносит в цепи ДНК двухцепочечный разрез со смещением. Как изображено на РИС. 19.38, разрез порождает одноцепочечные выступающие концы (по четыре нуклеотида). Нуклеаза HO не атакует мутантные локусы MAT, которые не поддаются переключению. Делеционный анализ показал, что последовательность в 24 пн (во всяком случае, бóльшая ее часть), покрывающая границу *Y*, необходима HO для расщепления *in vitro*. Сайт узнавания, довольно-таки крупный для эндонуклеазы, во всем геноме встречается лишь в трех кассетах по локусам *MAT*, *HML* и *HMR*.

Активной мишенью для эндонуклеазы служит только локус *MAT*, но не *HMR* с *HML*. Можно допустить, что те же механизмы, которые препятствуют транскрипции молчащих кассет, делают их недоступными и для действия эндонуклеазы HO. Эта неуязвимость гарантирует однонаправленность переключения.

Реакция, запускаемая HO-разрезанием, схематично — в рамках реакции между донорным и реципиентным участками — проиллюстрирована на РИС. 19.39. По характеру взаимодействий между отдельно взятыми цепями ДНК процесс вписывается в схему рекомбинации с внесением двухцепочечного разрыва (см. рис. 19.9). Этапы, следующие за исходным разрезом, задействуют те же ферменты, что и при гомологичной рекомбинации, а мутации по ряду соответствующих генов препятствуют переключению типа спаривания.

Предположим, свободный конец ДНК из локуса *MAT* вторгается в локус *HML* либо *HMR* и взаимодействует с гомологичной областью на его правой стороне. В это время область *Y* локуса *MAT* подвергается разрушению до момента, когда процесс достигает участка, гомологичного левой стороне *HML* (или *HMR*). На этот момент локус *MAT* спарен с *HML* (или с *HMR*), обоими

своими краями. В качестве замены участка, потерянного локусом *MAT*, копируется область *HML* (или *HMR*), содержащая в себе область *Y* (как минимум собственно *Y*), после чего спаренные локусы разделяются (порядок событий может быть и другим).

Как и в модели с двухцепочечным разрывом при рекомбинации, процесс переключения инициируется локусом (*MAT*), который и будет замещен. В этом смысле описание *HML* и *HMR* как донорных локусов относится к их предназначению, но не к механизму процесса. Подобно случаю репликативной транспозиции, сайт донора пребывает нетронутым, в то время как ДНК-реципиент претерпевает смену последовательности. Отличие от транспозиции — в том, что локус-реципиент переживает скорее замену, нежели прибавление, материала.

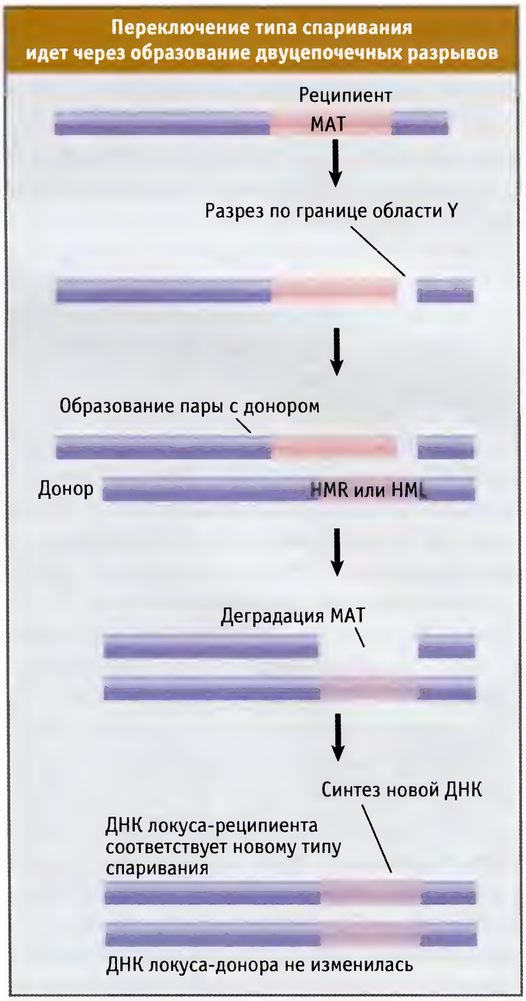


РИС. 19.39 Смена кассеты инициируется двухцепочечным разрывом ДНК в локусе-реципиенте (*MAT*) и сопровождается взаимодействием ДНК донора и реципиента в области Y и ее окрестностях

19.24 Переключение контролируется регуляцией экспрессии HO

Ключевые понятия:

- **Эндонуклеаза HO синтезируется в гаплоидных материнских клетках, и переключение «выливается» в новый тип спаривания обеих дочерних клеток.**

Продукция эндонуклеазы HO регулируется на уровне транскрипции. Существует три отдельных системы ее контроля:

- *HO* пребывает под контролем сообразно типу спаривания. Она не синтезируется в диплоидных клетках *MAT***a**/*MAT*α ввиду отсутствия потребности в переключении, когда оба аллеля *MAT* экспрессируются.
- Если *HO* экспрессируется в материнской, то не экспрессируется в дочерних клетках.
- Транскрипция *HO* зависит и от фазы клеточного цикла. Ген экспрессируется у материнской клетки только в конце фазы G1.

Такой график продукции нуклеазы по времени объясняет взаимосвязь переключения типа спаривания с последовательностью клеточных поколений. РИС. 19.40 показывает, как переключение детектируется только у продуктов деления; обе дочерние клетки имеют один и тот же тип спаривания, отличный (переключенный) от родительского. Причина в том, что привязка экспрессии *HO* к фазе G1 гарантирует переключение типа спаривания *до* репликации локуса *MAT*, в результате чего обоим потомкам достается новый тип спаривания.

Сайты, контролирующие транскрипцию *HO*, расположены на 1500 пн выше гена и действуют в *цис*-манере. Общий паттерн контроля таков, что репрессия по любому из многочисленных сайтов (как и ответ по нескольким регуляторным контурам) может предотвратить транскипцию *HO*. Классификация сайтов-участников представлена на РИС. 19.41.

Контроль типа спаривания напоминает таковой для прочих гаплоид-специфичных генов. Транскрипция предотвращается (в диплоидных клетках) репрессором **a**1/α2. Для этого репрессора на лежащем выше по ходу участке существует десять сайтов связывания. Последовательности этих сайтов различаются степенью соответствия консенсусу; неизвестно, сколько из них, и какие именно, требуются для гаплоид-специфичной репрессии.

Контроль транскрипции *HO* определяется «взаимной игрой» серий активирующих и репрессирующих событий. Гены *SWI1-5* требуются для транскипции *HO*. Их функция — предотвращать репрессию *HO* продуктами генов *SIN1-6*. Сначала были открыты гены *SWI* (у мутантов, неспособных к переключению); затем были открыты гены *SIN* (благодаря своей способности снимать блок, вызванный конкретными мутациями *SWI*). Взаимовлияние белков Swi и Sin привязывает экспрес-

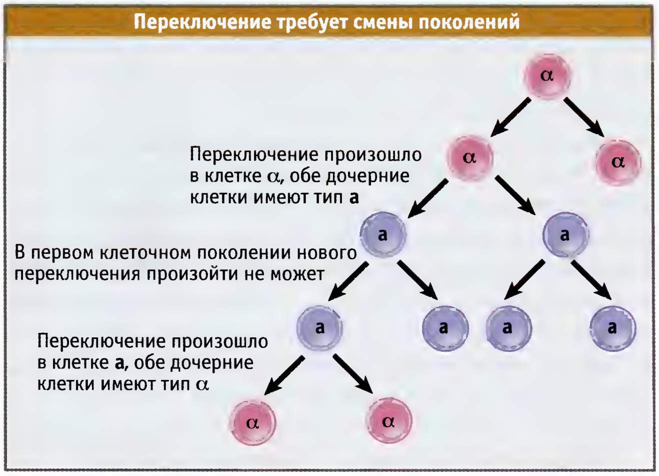


РИС. 19.40 Если в какой-нибудь клетке произошло переключение типа спаривания, это проявится у обеих дочерних клеток. Прежде чем снова переключить тип спаривания, дочерним клеткам надо еще раз поделиться (то есть пройти полный клеточный цикл)

сию *HO* к определенному временному «окну» в клеточном цикле ряда поколений дрожжевых клеток.

Некоторые из генов *SWI* и *SIN* (вернее, их белки) не специализированы на типе спаривания, но являются глобальными регуляторами транскрипции, чьи функции нужны для экспрессии многих локусов. К таким относятся перестраивающий хроматин комплекс *SWI/SNF*, в состав которого входят Swi1,2,3, а также

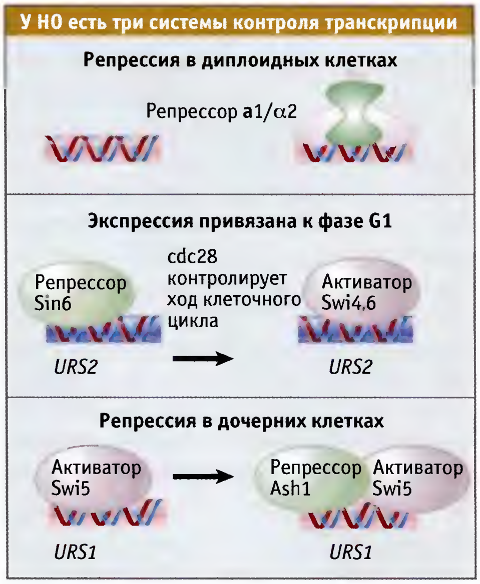


РИС. 19.41 Транскрипция гена *HO* находится под влиянием трех регуляторных систем и происходит только с разрешения всех трех

продукты локусов *SIN1-4*, кодирующих белки-регуляторы состояний хроматина. Их участие в регуляции типа спаривания — побочное. Истинным регулятором, таким образом, является белок Swi, специфически противодействующий общей системе репрессии в локусе *HO*.

Контроль соответствия клеточному циклу наводит локус *URS2*, содержащий девять копий октануклеотидной последовательности. Копия консенсуса, будучи прикреплена к любому гену, может привести его экспрессию к такому же соответствию. Со своей стороны, ген, подшитый к такой последовательности, репрессирован всегда, за исключением непродолжительного периода ближе к концу фазы G1. Активаторами, отменяющими репрессию *HO* по *URS2*, служат Swi4 и Swi6. Их активность зависит от функции регулятора клеточного цикла Cdc28, «олицетворяющего» обреченность клетки на скорое деление.

Передаточным звеном привязки экспрессии *HO* к чередующимся поколениям служит активатор Swi5, противодействующий системе общей репрессии в исполнении белков Sin3,4. Эта система репрессии действует через элементы *URS1*, расположенные выше, чем от *HO*. У мутантов, лишенных функций Sin3,4, *HO* одинаково хорошо транскрибируется в двух клеточных поколениях подряд.

Сам по себе Swi5 не специфицирует материнскую клетку по отношению к дочерней. Его эффект обусловлен противостоянием с Ash1p — репрессором, накапливающимся в дочерних клетках в конце анафазы. Мутации по *ASH1* позволяют дочерним клеткам, подобно материнской, переключать тип спаривания. Локализация Ash1p обусловлена транспортом его мРНК из материнской клетки вдоль актиновых филаментов в почку-дочку (см. разд. 7.16 «Местонахождение мРНК может быть четко детерминировано»). Его присутствие удерживает Swi5 от активации гена *HO*. Белок Ash1p действует, связываясь с многочисленными копиями консенсусной последовательности, распределенными по регуляторным областям *URS1* и *URS2*. Когда дочерние клетки вырастают, чтобы снова поделиться, Ash1p в них становится разбавлен до концентрации, позволяющей гену *HO* снова начать экспрессироваться.

19.25 Резюме

Под рекомбинацией подразумевают физический обмен частями между соответствующими молекулами ДНК. В результате получается дуплекс ДНК, у которого два участка, исходно принадлежавшие разным родительским дуплексам, соединяются фрагментом гибридной (гетеродуплексной) ДНК, ведущей свое происхождение от разных родителей, по одной цепи от каждого. Если гибридная ДНК содержит неспаренные основания, то может иметь место их коррекция.

Иногда образование гибридной ДНК не сопровождается рекомбинацией между маркерами по разные ее стороны, то есть рекомбинации между последовательностями флангов не происходит. Конверсия гена происходит, если в ходе нормальной рекомбинации образуется протяженный участок гибридной ДНК, который подвергается коррекции в соответствии с *одной* из родительских цепей, причем один из генов, по крайней мере частично, преобразуется в последовательность другого. Конверсия между неаллельными генами также возможна — как аномальное событие.

Рекомбинацию инициирует разрыв двух цепей дуплекса ДНК. Разрыв разрастается в пробел с выступающим одноцепочечным краем; этот свободный конец образует гетеродуплекс с аллельной последовательностью. ДНК, подвергшаяся инициирующему разрыву, затем подвергается частичной замене себя на последовательность хромосомы, в которую она вторглась (почему инициирующую ДНК и называют реципиентом). Горячие точки рекомбинации — это сайты, по которым инициируются двухцепочечные разрывы. Градиент конверсии гена отражает распределение вероятности преобразования материала ДНК в свободный одноцепочечный конец, которая убывает по мере удаления от разрыва (горячей точки).

У дрожжей рекомбинация инициируется белком Spo11. Это топоизомеразо-подобный фермент, связывающийся со свободным 5′-концом ДНК. После этого DSB продолжает генерировать одноцепочечную ДНК, которая может взаимодействовать со своим комплементом на другой хромосоме. Мутации дрожжей, блокирующие образование в их клетках синаптонемного комплекса, показывают необходимость рекомбинации для его образования. Возможно, образование синаптонемного комплекса инициируется двухцепочечными разрывами. Он может продолжать свое существование до полного завершения рекомбинации. Мутации, поражающие компоненты синаптонемного комплекса, блокируют его образование, но не препятствуют попарной расстановке хромосом, так что взаимное распознавание гомологов не зависит от рекомбинации и образования синаптонемного комплекса.

Полный набор реакций, требуемый для рекомбинации, может быть осуществлен белками Rec и Ruv *E. coli*. Одноцепочечный участок со свободным концом порождается нуклеазой RecBCD. Фермент связывается с ДНК по одну сторону последовательности *chi*, а затем движется по направлению к *chi*, попутно расплетая ДНК. В последовательность *chi* вносится одноцепочечный разрыв. Последовательности *chi* служат горячими точками рекомбинации. Одиночная цепь ДНК служит субстратом для RecA, который обладает способностью сближать гомологичные молекулы. Этот белок поддерживает реакцию вторжения одиночной цепи из одной молекулы в дуплекс другой молекулы. Гетеродуплекс ДНК образуется через замещение одной из цепей исходного дуплекса. Так образуется рекомбинантное соединение, разрезание которого осуществляют белки Ruv. На гетеродуплексном участке работают RuvA и RuvB, в то время как RuvC занимается собственно разделением, то есть разрезает структуру Холлидея.

Рекомбинация, подобно репликации и (возможно) транскрипции, требует топологических манипуляций с ДНК. Топоизомеразы могут расслаблять (или вводить новые) сверхвитки в ДНК и требуются для распутывания молекул, конкатенированных как следствие рекомбинации или репликации. Топоизомеразы первого типа вносят разрыв в одну цепь дуплекса ДНК; топоизомеразы второго типа разрывают обе. Фермент образует сшивку с ДНК в виде химической связи с 5′-фосфатом (по типу А) или с 3′-фосфатом (по типу B).

Ферменты-участники сайт-специфичной рекомбинации по механизму действия родственны топоизомеразам. Под именем «рекомбиназы» их объединяют в общий класс; к отдельному подклассу интеграз относятся рекомбиназы, задействованные в интеграции фага (опосредующие взаимодействие ДНК фага с хромосомой бактерии). Система Cre/*lox* использует по две молекулы Cre для связывания с каждым из сайтов *lox*, так что рекомбиназный комплекс представляет собой тетрамер. Эта система — один из стандартных инструментов внедрения ДНК в чужеродный геном. Интеграция фага лямбда требует наличия белка Int фага и белка IHF хозяина (*E. coli*). Она основана на прицельном разрыве и воссоединении цепей при отсутствии какого-либо синтеза ДНК. Реакция использует упаковку *attP*-фрагмента фаговой ДНК в нуклеопротеиновую структуру интасомы, содержащую по нескольку копий Int и IHF; со всем эти массивом связывается последовательность *attB* хозяина, что ведет к рекомбинации. Реакция в обратном направлении требует присутствия фагового белка Xis. Некоторые интегразы действуют в *цис*-манере, при которой тирозин, образующий сшивку с полусайтом ДНК, предоставляется субъединицей фермента, с этим же полусайтом и связанной. Другие интегразы производят *транс*-разрез, при котором свой тирозин предоставляет другая субъединица.

Дрожжи *S. cerevisiae* могут распространяться как в гаплоидном, так и в диплоидном состоянии. Взаимные превращения этих состояний происходят при спаривании (слияние гаплоидных клеток с образованием диплоидной) либо при споруляции (мейоз диплоидных клеток с образованием гаплоидных спор). Способность к участию в этих мероприятиях определяется типом спаривания, присущим штамму. Тип спаривания определяется последовательностью ДНК в локусе *MAT*. Посредством акта рекомбинации, ведущего к замене этой последовательности, тип спаривания может быть переключен на противоположный. Акт рекомбинации инициируется двухцепочечным разрывом — подобно акту гомологичной рекомбинации — однако последующий ход событий гарантирует однонаправленное замещение последовательности в локусе *MAT*.

Замещение регулируется таким образом, что *MAT*α, как правило, замещается последовательностью из *HMR***a**, в то время как *MAT***a** обычно замещается последовательностью из *HML*α. Реакцию запускает эндонуклеаза HO, которая опознает в локусе *MAT* уникальную последовательность-мишень. Эндонуклеаза HO регулируется на уровне транскрипции системой, которая гарантирует ее экспрессию в материнских, но не дочерних, клетках. Как следствие, обоим потомкам присущ одинаковый — новый — тип спаривания.

Литература

19.4 Двухцепочечные разрывы инициируют рекомбинацию

Исследование

Hunter, N. and Kleckner, N. (2001). The single-end invasion: an asymmetric intermediate at the double-strand break to double-Holliday junction transition of meiotic recombination. *Cell* 106, 59–70.

Обзор

Lichten, M. and Goldman, A. S. (1995). Meiotic recombination hotspots. *Annu. Rev. Genet.* 29, 423–444.

Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J., and Stahl, F. W. (1983). The double-strand-break repair model for recombination. *Cell* 33, 25–35.

19.5 Рекомбинирующие хромосомы соединены синаптонемным комплексом

Исследование

Blat, Y, and Kleckner, N. (1999). Cohesins bind to preferential sites along yeast chromosome III, with differential regulation along arms versus the central region. *Cell* 98, 249–259.

Dong, H. and Roeder, G. S. (2000). Organization of the yeast Zip1 protein within the central region of the synaptonemal complex. *J. Cell Biol.* 148, 417–426.

Klein, F. et al. (1999). A central role for cohesins in sister chromatid cohesion, formation of axial elements, and recombination during yeast meiosis. *Cell* 98, 91–103.

Sym, M., Engebrecht, J. A., and Roeder, G. S. (1993). ZIP1 is a synaptonemal complex protein required for meiotic chromosome synapsis. *Cell* 72, 365–378.

Обзор

Roeder, G. S. (1997). Meiotic chromosomes: it takes two to tango. *Genes Dev.* 11, 2600–2621.

Zickler, D. and Kleckner, N. (1999). Meiotic chromosomes: integrating structure and function. *Annu. Rev. Genet.* 33, 603–754.

19.6 Синаптонемный комплекс образуется после внесения двухцепочечных разрывов

Исследование

Allers, T. and Lichten, M. (2001). Differential timing and control of noncrossover and crossover recombination during meiosis. *Cell* 106, 47–57.

Weiner, B. M. and Kleckner, N. (1994). Chromosome pairing via multiple interstitial interactions before and during meiosis in yeast. *Cell* 77, 977–991.

Обзор

McKim, K. S., Jang, J. K., and Manheim, E. A. (2002). Meiotic recombination and chromosome segregation in *Drosophila* females. *Annu. Rev. Genet.* 36, 205–232.

Petes, T. D. (2001). Meiotic recombination hot spots and cold spots. *Nat. Rev. Genet.* 2, 360–369.

19.8 Бактериальная система RecBCD стимулируется последовательностями *chi*

Исследование

Dillingham, M. S., Spies, M., and Kowalczykowski, S. C. (2003). RecBCD enzyme is a bipolar DNA helicase. *Nature* 423, 893–897.

Spies, M., Bianco, P. R., Dillingham, M. S., Handa, N., Baskin, R. J., and Kowalczykowski, S. C. (2003). A molecular throttle: the recombination hotspot chi controls DNA translocation by the RecBCD helicase. *Cell* 114, 647–654.

Taylor, A. F. and Smith, G. R. (2003). RecBCD enzyme is a DNA helicase with fast and slow motors of opposite polarity. *Nature* 423, 889–893.

19.9 Белки переноса цепи катализируют ассимиляцию одноцепочечной ДНК

Обзор

Kowalczykowski, S. C., Dixon, D. A., Eggleston, A. K., Lauder, S. D., and Rehrauer, W. M. (1994). Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol. Rev.* 58, 401–465.

Kowalczykowski, S. C. and Eggleston, A. K. (1994). Homologous pairing and DNA strand-exchange proteins. *Annu. Rev. Biochem.* 63, 991–1043.

Lusetti, S. L. and Cox, M. M. (2002). The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. *Annu. Rev. Biochem.* 71, 71–100.

19.10 Система Ruv разделяет структуры Холлидея

Исследование

Boddy, M. N., Gaillard, P. H., McDonald, W. H., Shanahan, P., Yates, J. R., and Russell, P. (2001). Mus81-Eme1 are essential components of a Holliday junction resolvase. *Cell* 107, 537–548.

Chen, X. B., Melchionna, R., Denis, C. M., Gaillard, P. H., Blasina, A., Van de Weyer, I., Boddy, M. N., Russell, P., Vialard, J., and McGowan, C. H. (2001). Human Mus81-associated endonuclease cleaves Holliday junctions *in vitro*. *Mol. Cell* 8, 1117–1127.

Constantinou, A., Davies, A. A., and West, S. C. (2001). Branch migration and Holliday junction resolution catalyzed by activities from mammalian cells. *Cell* 104, 259–268.

Kaliraman, V., Mullen, J. R., Fricke, W. M., Bastin-Shanower, S. A., and Brill, S. J. (2001). Functional overlap between Sgs1-Top3 and the Mms4-Mus81 endonuclease. *Genes Dev.* 15, 2730–2740.

Обзор

Lilley, D. M. and White, M. F. (2001). The junction-resolving enzymes. *Nat. Rev. Mol. Cell Biol.* 2, 433–443.

West, S. C. (1997). Processing of recombination intermediates by the RuvABC proteins. *Annu. Rev. Genet.* 31, 213–244.

19.13 Топоизомеразы изменяют число сверхвитков, релаксируя или сверхспирализуя ДНК

Обзор

Champoux, J. J. (2001). DNA topoisomerases: structure, function, and mechanism. *Annu. Rev. Biochem.* 70, 369–413.

Wang, J. C. (2002). Cellular roles of DNA topoisomerases: a molecular perspective. *Nat. Rev. Mol. Cell Biol.* 3, 430–440.

19.14 Топоизомеразы разрывают цепи ДНК и снова их склеивают

Исследование

Lima, C. D., Wang, J. C., and Mondragon, A. (1994). Three-dimensional structure of the 67K N-terminal fragment of *E. coli* DNA topoisomerase I. *Nature* 367, 138–146.

Обзор

Champoux, J. J. (2001). DNA topoisomerases: structure, function, and mechanism. *Annu. Rev. Biochem.* 70, 369–413.

19.16 Специализированная рекомбинация задействует специфичные сайты

Исследование

Metzger, D., Clifford, J., Chiba, H., and Chambon, P. (1995). Conditional site-specific recombination in mammalian cells using a ligand-dependent chimeric Cre recombinase. *Proc. Natl. Acad. Sci. USA* 92, 6991–6995.

Nunes-Duby, S. E., Kwon, H. J., Tirumalai, R. S., Ellenberger, T., and Landy, A. (1998). Similarities and differences among 105 members of the Int family of site-specific recombinases. *Nucleic Acids Res.* 26, 391–406.

Обзор

Craig, N. L. (1988). The mechanism of conservative site-specific recombination. *Annu. Rev. Genet.* 22, 77–105.

19.18 Сайт-специфичная рекомбинация напоминает топоизомеразную активность

Исследование

Guo, F., Gopaul, D. N., and van Duyne, G. D. (1997). Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature* 389, 40–46.

19.19 Рекомбинация фага лямбда происходит в интасоме

Исследование

Biswas, T., Aihara, H., Radman-Livaja, M., Filman, D., Landy, A., and Ellenberger, T. (2005). A structural basis for allosteric control of DNA recombination by lambda integrase. *Nature* 435, 1059–1066.

Wojciak, J. M., Sarkar, D., Landy, A., and Clubb, R. T. (2002). Arm-site binding by lambda integrase: solution structure and functional characterization of its amino-terminal domain. *Proc. Natl. Acad. Sci. USA* 99, 3434–3439.

Системы репарации

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

20.1 Введение

20.2 Системы репарации корректируют повреждения ДНК

- Системы репарации распознают последовательности ДНК, не соответствующие стандартам спаривания оснований.
- Системы эксцизионной репарации удаляют одну цепь ДНК по месту повреждения, а затем синтезируют ее.
- Системы рекомбинационной репарации используют рекомбинацию для замены поврежденного участка двойной цепи.
- Все эти системы предрасположены к внесению ошибок в последовательность ДНК.
- Фотореактивация представляет собой немутагенную систему репарации, специфично действующую на пиримидиновые димеры.

20.3 Системы эксцизионной репарации *E. coli*

- Система Uvr вносит надсечки в поврежденную цепь ДНК с промежутком ~12 пн, содержащим собственно повреждение; затем система удаляет ДНК между надсечками и синтезирует новую.

20.4 Способы эксцизионной репарации в клетках млекопитающих

- У млекопитающих запуск эксцизионной репарации производится прямым удалением поврежденного основания ДНК.
- Удаление основания провоцирует удаление и замену нуклеотидного отрезка.
- Какой из двух способов эксцизионной репарации будет активирован, определяется природой реакции удаления основания.
- Путь polδ/ε приводит к замене полинуклеотидного отрезка; путь polβ приводит к замене одного нуклеотида.

20.5 Метилазы и гликозилазы вынимают основания из ДНК

- Гликозилазы узнают и напрямую удаляют из ДНК урацил и алкилированные основания.
- Исправление пиримидиновых димеров производится разрывом ковалентной связи между ними.
- Метилазы добавляют к цитозину метильную группу.
- Все эти ферменты действуют с выведением азотистого основания за пределы двойной спирали, там они удаляют, либо модифицируют его и возвращают в спираль.

20.6 Ошибки репаративного синтеза и фенотипы мутаторов

- Во время репликации поврежденная и нерепарированная ДНК приводит к застреванию ДНК-полимеразы III.
- Цепь, комплементарную поврежденной, могут синтезировать ДНК-полимераза V (кодируемая *umuCD*) или ДНК-полимераза IV (кодируемая *dinB*).
- Последовательность ДНК, синтезированной репаративной ДНК-полимеразой, часто содержит ошибки.
- Гены-мутаторы кодируют белки, от которых зависит точность репликации; мутации мутаторов повышают уровень спонтанного мутагенеза.

20.7 Контроль направления репарации несовершенных пар оснований

- Гены *mut* кодируют систему репарации некомплементарностей по отдельным парам оснований.
- Выбор, в какой из цепей произвести замену нуклеотида для восстановления комплементарности, не случаен.
- В полуметилированном $\begin{matrix}\text{GATC}\\ \text{CTAG}\end{matrix}$ обычно заменяется неметилированная цепь.
- Эта система репарации используется для удаления ошибок в новосинтезированной цепи ДНК. В случае неспаренных G-T или C-T предпочтительно удаляется T.

20.8 Системы рекомбинационной репарации *E. coli*

- Гены *rec E. coli* кодируют компоненты системы заимствования информации из другого дуплекса.
- Система заимствования информации из другого дуплекса приходит в действие, когда репликация оставляет в новой цепи ДНК брешь напротив поврежденного участка матрицы.
- Чтобы восполнить пробел, используется одиночная цепь другого дуплекса.
- Затем поврежденная последовательность удаляется и синтезируется заново.

20.9 Рекомбинация служит важным механизмом исправления ошибок репликации

- Репликативная вилка может застрять при встрече в ДНК поврежденного участка или ника.
- Репарация застрявшей вилки может включать в себя спаривание новосинтезированных цепей.

- Застрявшая вилка может заново стартовать после репарации повреждения, используя для продвижения хеликазу.
- Застрявшая вилка представляет собой структуру Холлидея и может быть разведена на дуплексы разрывом двух цепей при участии резольваз.

20.1 Введение

Любой случай внесения в ДНК отклонений от обыкновенной двуспиральной структуры представляет угрозу для генетической конституции клетки. Вред, наносимый ДНК, минимизируют системы, распознающие повреждения и корректирующие их. Системы репарации* столь же сложны, сколь и сам по себе репликативный аппарат, что показывает их важность для выживания клетки. Когда репаративная система исправляет изменение ДНК, оно не имеет последствий. Если системе это не удается, результатом может быть мутация. Измеряемая скорость возникновения мутаций отражает баланс между накоплением повреждений ДНК и скоростью их исправления (или попыток исправления).

Многие системы репарации умеют воспринимать как сигнал к действию целый спектр искажений структуры ДНК, а клетки склонны иметь по нескольку систем, способных справляться с повреждениями ДНК. Для эукариот важность репарации огромна; в геноме человека идентифицировано >130 генов, причастных к репарационным процессам. Как показано на РИС. 20.1, системы репарации можно подразделить на несколько общих типов:

- Некоторые ферменты напрямую исправляют конкретные разновидности повреждений ДНК.
- Существуют пути репарации с вырезанием азотистых оснований или целых нуклеотидов. Все они работают по принципу удаления и замены материала.
- Существуют системы, использующие рекомбинацию для привлечения нетронутой копии последователь-

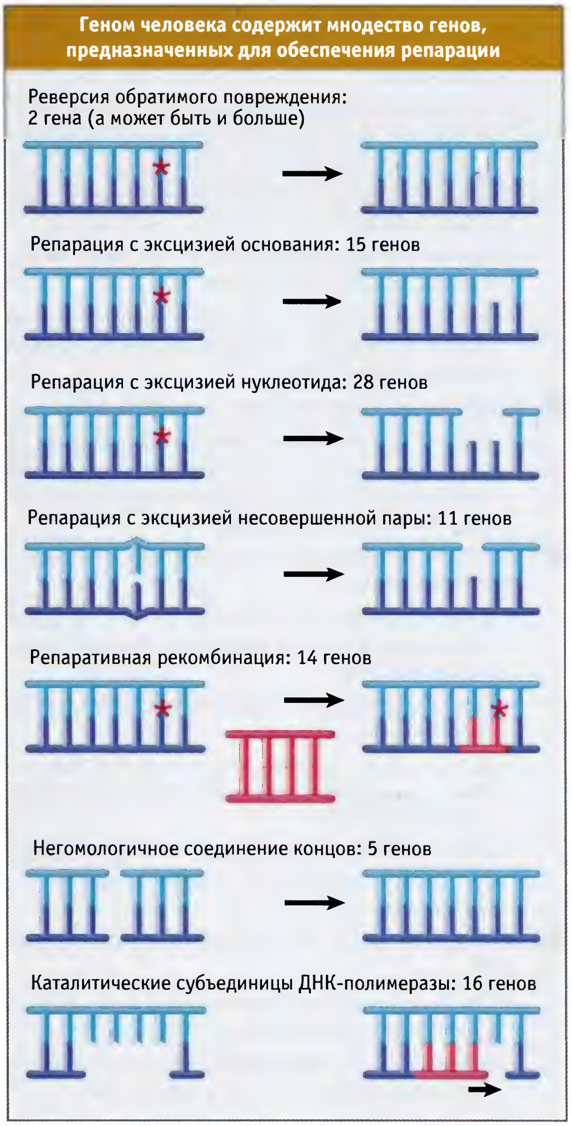


РИС. 20.1 Гены, имеющие отношение к репарации, делят на группы по участию в различных путях преодоления повреждений в ДНК

* Англ. *repair* — починка, восстановление, ремонт. В слове repair, применительно к двойной спирали комплементарных цепей ДНК, можно разглядеть и другой смысл: "re-pair", «восстановление парности». — *Прим. перев.*

ности, которая используется для замещения поврежденного участка дуплекса.

- Путь негомологичного соединения концов склеивает двухцепочечные «рваные» концы.
- Синтезом отрезков ДНК для замещения может заниматься несколько разных ДНК-полимераз.

Прямая репарация, при которой повреждение просто удаляется или исправляется на месте, встречается редко. Хорошим примером служит *фотореактивация* пиримидиновых димеров, при которой проблемные ковалентные связи восстанавливаются светозависимым ферментом. Фотореактивация широко распространена в живых системах (хотя отсутствует у плацентарных млекопитающих) и, по-видимому, особенно важна для растений. У *Escherichia coli* ее работа зависит от единственного гена *phr*, кодирующего фермент фотолиазу.

Несовершенность пар оснований представляет одну из главных целей для систем репарации. Для *репарации некомплементарностей* ДНК тщательно обыскивается на предмет азотистых оснований, которые, располагаясь в разных цепях ДНК друг напротив друга, не образуют правильную пару. Некомплементарности, возникающие в процессе репликации, корректируются с различением новой и старой цепей и внесением исправлений преимущественно в новосинтезированный материал. Другие системы работают с последствиями химической модификации оснований, таких как дезаминирование. Важность этих систем подчеркивает тот факт, что мутации в генах, родственных дрожжевым генам, причастным к репарации некомплементарностей, вызывают рак у людей.

Неспаренные основания обычно корректируются *эксцизионной репарацией*. Ее инициирует фермент опознания, который обнаруживает или само поврежденное основание, или изменение в пространственной конфигурации ДНК. Системы эксцизионной репарации бывают двух типов:

- При *репарации с эксцизией оснований* система напрямую вырезает из ДНК поврежденное азотистое основание и заменяет его. Хорошим примером служит урацил-гликозилаза, удаляющая урацилы, некомплементарные гуанинам (см. разд. 20.5 «Метилазы и гликозилазы вынимают основания из ДНК»).
- При *репарации с эксцизией нуклеотидов* система вырезает из ДНК последовательность, содержащую поврежденное основание (или несколько поврежденных оснований); затем на замену вырезанному материалу синтезируется новый отрезок ДНК.

Основные рабочие моменты эксцизионной репарации резюмирует РИС. 20.2. Подобные системы есть у всех организмов. Некоторые из них распознают в ДНК любые повреждения; другие настроены на отдельные типы поврежденных оснований. Нередко на один клеточный тип приходится множество различных систем эксцизионной репарации.

Системы *рекомбинационной репарации* имеют дело с ситуациями, при которых матричная цепь ДНК оказалась поврежденной, а репликация вынуждена была «проскочить» это место, оставив за собой брешь в дочерней цепи. Рекомбинация используется, чтобы привлечь к восстановлению поврежденной последовательности ее интактную копию из другого (гомологичного) дуплекса и заполнить брешь с помощью этой копии.

Отличительной особенностью рекомбинационной репарации является необходимость иметь дело с двухцепочечными разрывами (ДЦР). ДЦР инициируют переходы цепей между дуплексами ДНК при гомологичной рекомбинации. Также они могут возникать при проблемах с репликацией и в этом случае срабатывать на запуск систем рекомбинационной репарации. Если ДЦР образуются как повреждение вследствие влияний окружающей среды (например, при облучении) или являются результатом укорочения теломер, они могут вызвать мутации. Одна из систем обращения с ДЦР склеивает негомологичные концы ДНК.

Мутации, нарушающие способность клетки *E. coli* репарировать свою ДНК, распадаются на группы в соответствии с несколькими путями репарации (необязательно независимыми между собой). Наиболее известными путями являются система эксцизионной репарации *uvr*, метил-зависимая система репарации некомплементарностей, а также *recB*- и *recF*-пути рекомбинации и рекомбинационной репарации. С этими системами ассоциированы ферментные активности эндо- и экзонуклеаз (важных для удаления поврежденной ДНК), резольваз (эндонуклеаз со специфическим действием в рекомбинантных соединениях), расплетающих ДНК хеликаз, и ДНК-полимераз, синтезирующих новую ДНК. Некоторые из этих активностей уникальны для конкретного пути репарации, другие же участвуют во множестве путей.

Аппарат репликации уделяет огромное внимание контролю качества. ДНК-полимеразы, используя

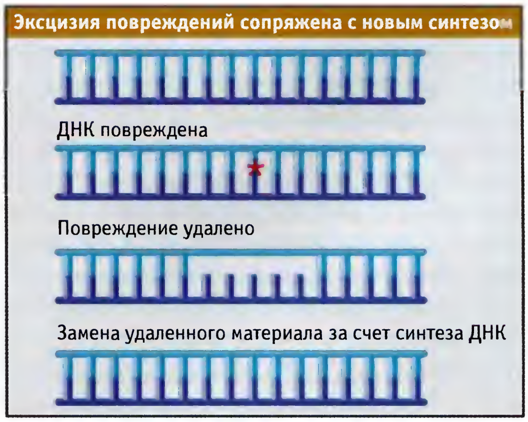


РИС. 20.2 При эксцизионной репарации брешь, возникшая на месте вырезанного повреждения, нуждается в «заплатке»

собственную корректирующую активность, выверяют последовательность дочерней цепи и устраняют ошибки. Некоторые из систем репарации менее аккуратны, когда синтезируют ДНК для замены поврежденного материала. За это их назвали системами *«склонными ошибаться»*.

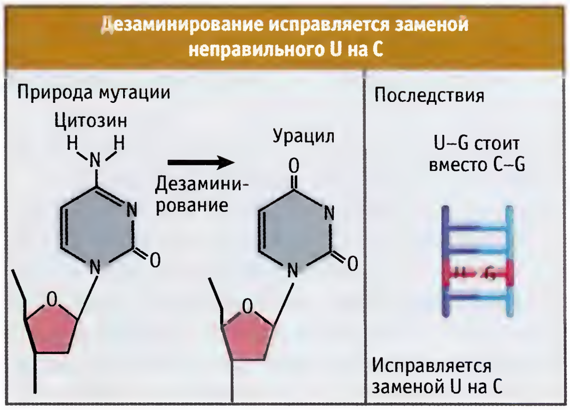
20.2 Системы репарации корректируют повреждения ДНК

Ключевые понятия:

- **Системы репарации распознают последовательности ДНК, не соответствующие стандартам спаривания оснований.**
- **Системы эксцизионной репарации удаляют одну цепь ДНК по месту повреждения, а затем синтезируют ее.**
- **Системы рекомбинационной репарации используют рекомбинацию для замены поврежденного участка двойной цепи.**
- **Все эти системы предрасположены к внесению ошибок в последовательность ДНК.**
- **Фотореактивация представляет собой немутагенную систему репарации, специфично действующую на пиримидиновые димеры.**

Типы повреждений, ведущих к запуску систем репарации, можно, в общем, подразделить на два класса:

- При *однонуклеотидных заменах* нарушенной оказывается нуклеотидная последовательность, но не структура ДНК в целом. Однонуклеотидные замены не влияют на процессы транскрипции или репликации (при которых цепи дуплекса разделены), но имеют эффект у будущих поколений клеток. Нарушение можно рассматривать как преобразование одного из азотистых оснований в другое, теперь уже плохо спаренное со своим партнером в комплементарной цепи. Однонуклеотидная замена может возникнуть в результате нарушения основания *in situ* или же быть следствием ошибки репликации. На РИС. 20.3 показано, как дезаминирование цитозина до урацила (спонтанно или под воздействием вещества-мутагена) порождает некомплементарную пару U-G. На РИС. 20.4 показано, как в процессе репликации при ошибочном встраивании аденина вместо цитозина образуется пара A-G. Сходные последствия может иметь ковалентное присоединение к азотистому основанию небольшого радикала, модифицирующего способность основания образовывать пару с другим основанием. Эти изменения могут выражаться как в очень незначительном искажении (в случае пары U-G), так и во вполне ощутимом изменении структуры двойной спирали (в случае пары A-G). Однако общий момент состоит в том, что некомплементарность сохраняется только до следующей



Дезаминирование цитозина приводит к наличию в ДНК пары U-G. В этой паре замене подлежит, как правило, урацил

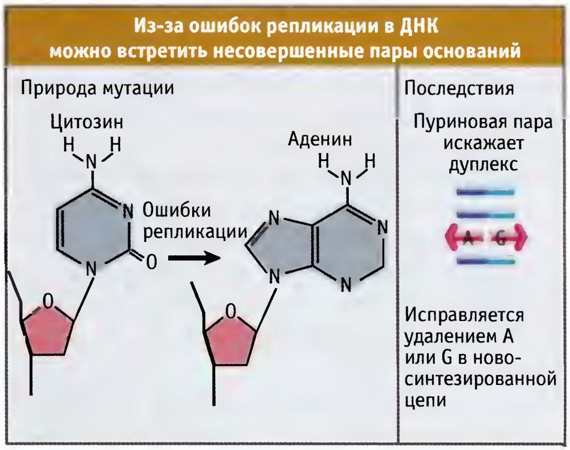


РИС. 20.4 Ошибка репликации приводит к наличию несовершенной пары оснований. Если ее не исправить, в следующем раунде репликации мутация будет закреплена

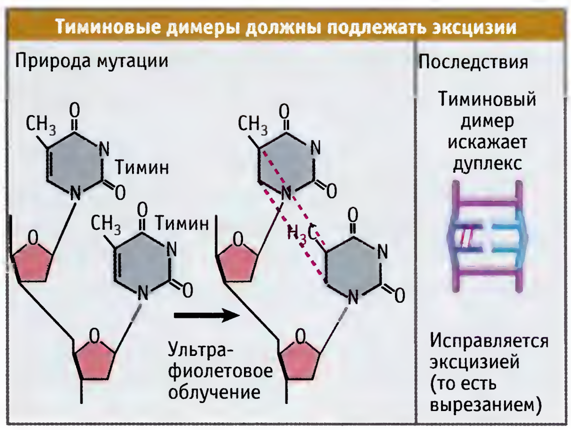


РИС. 20.5 Ультрафиолетовое облучение приводит к димеризации соседних тиминов. Наличие такого димера блокирует репликацию и транскрипцию

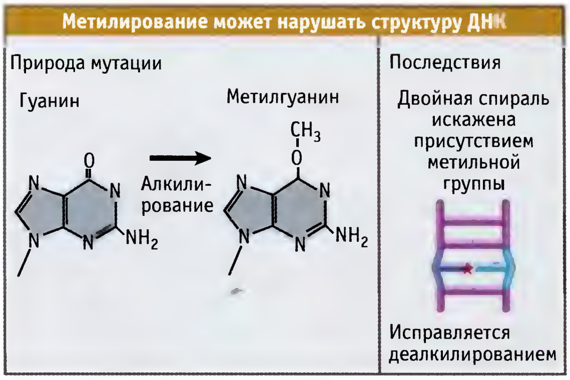


РИС. 20.6 Метилирование основания нарушает геометрию двойной спирали и приводит к ошибкам репликации. Метильная группа обозначена звездочкой

репликации. Таким образом время, отведенное на ремонт повреждения до того, как оно будет зафиксировано репликацией, ограничено.

- **Искажения структуры** ДНК могут создавать физические помехи при репликации или транскрипции. Внесение ковалентных сшивок между основаниями одной цепи ДНК или между основаниями противоположных цепей ингибирует репликацию и транскрипцию. На РИС. 20.5 приведен пример того, как ультрафиолетовое (УФ) облучение, сшив ковалентными связями два смежных тимина, привело к образованию внутрицепочечного **пиримидинового димера**. РИС. 20.6 показывает, как добавление к одному из оснований громоздкого придатка деформирует двойную спираль. Разрыв или удаленное основание одной из цепей ДНК не позволит этой цепи быть матрицей, подходящей для синтеза РНК или ДНК (это показано на РИС. 20.7). Общность всех этих изменений состоит в том, что вредоносная химическая модификация будет оставаться в ДНК и вызывать структурные

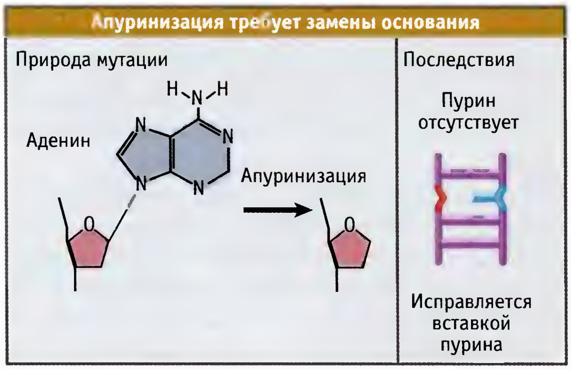


РИС. 20.7 Апуринизация заключается в удалении азотистого основания (как правило — пурина), что приводит к блоку репликации и транскрипции

дефекты и/или провоцировать мутации до тех пор, пока не будет удалена.

Если системы репарации выключить, клетка становится чрезвычайно чувствительной к ультрафиолетовому облучению. Исторически внесение УФ-индуцированных повреждений служило основным способом тестирования систем репарации, и, проводя на основании таких тестов сравнительную оценку их активностей и относительных эффективностей, следует иметь в виду, что при использовании других методик внесения повреждений в ДНК акценты (касательно характера повреждений и способов борьбы с ними) могут быть расставлены по другому.

20.3 Системы эксцизионной репарации *E. coli*

Ключевые понятия:

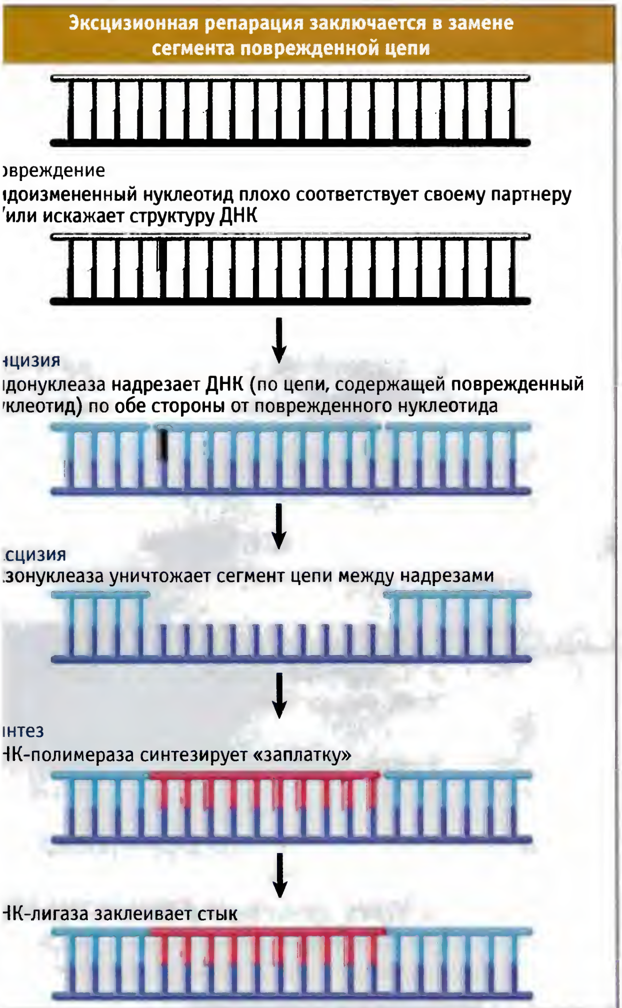
- Система Uvr вносит надсечки в поврежденную цепь ДНК с промежутком ~12 пн, содержащим собственно повреждение; затем система удаляет ДНК между надсечками и синтезирует новую.

Системы эксцизионной репарации варьируют по своей специфичности, но по своему действию подчиняются одним и тем же общим принципам. Каждая такая система удаляет из ДНК некомплементарные или поврежденные основания, а затем синтезирует новый отрезок ДНК, чтобы их заменить. Обобщенная последовательность событий проиллюстрирована на РИС. 20.8.

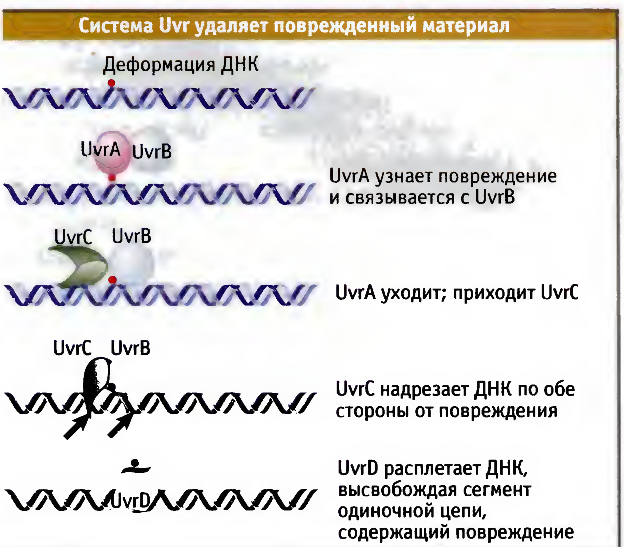
На этапе внесения надсечек (**инцизии**) поврежденная структура опознается эндонуклеазой, которая разрезает цепь ДНК по обе стороны от повреждения. На этапе вырезания (собственно, **эксцизии**) 5′–3′ экзонуклеаза удаляет отрезок поврежденной цепи.

На этапе *синтеза* образовавшийся одноцепочечный участок служит матрицей для ДНК-полимеразы, синтезирующей замену вырезанной последовательности. (Синтез нового материала может быть сопряжен с удалением старого в одно координированное действие.) По окончании синтеза ДНК-лигаза образует ковалентную сшивку между новым фрагментом и исходной ДНК.

Система эксцизионной репарации *uvr* состоит из трех генов, *uvrA*, *uvrB* и *uvrC*, кодирующих компоненты репаративной эндонуклеазы. Она функционирует по этапам, обозначенным на РИС. 20.9. Сначала комбинация UvrAB опознает пиримидиновые димеры и прочие громоздкие «патологии». Следующим этапом UvrA диссоциирует (для этого требуется аденозинтрифосфат, АТФ), а к UvrB подсоединяется UvrC. Комбинация UvrBC делает по надсечке с каждой стороны от повреждения: с отступом в семь нуклеотидов от 5′-края и с отступом в три-четыре нуклеотида от 3′-края. Для этого тоже требуется АТФ. Хеликаза UvrD помогает расплести ДНК между надсечками, чтобы высвободить поврежденную цепь, ко-



:. 20.8 При эксцизионной репарации происходит замена отрез-
цепи, где есть один или несколько поврежденных нуклеотидов



. 20.9 Система Uvr действует поэтапно: UvrA находит повреж-
ние, UvrBC надрезает ДНК, UvrD выплетает из ДНК испорченный
мент

торую уничтожает ДНК-полимераза I. Ферментом, участвующим в репаративном синтезе, похоже, является ДНК-полимераза I (хотя ее могут заменить ДНК-полимеразы II и III). В клетках эукариот аналогичный путь репарации состоит из тех же событий, но их порядок отличается (см. рис. 20.23).

Практически вся эксцизионная репарация в клетке *E. coli* происходит за счет UvrABC. Почти всегда (в 99% случаев) средняя длина замененного отрезка ДНК составляет 12 нуклеотидов. (В связи с этим ее иногда называют «ремонт короткими заплатками».) Оставшийся 1% случаев — это замена отрезков ДНК по большей части в ~1500 пн, хотя некоторые из них могут достигать >9000 пн (что иногда называют «ремонт длинными заплатками»). Неизвестно, почему какие-то события запускают режим «длинных заплаток» вместо режима «коротких».

Комплекс Uvr могжет быть направлен к поврежденным сайтам другими белками. Если повреждение ДНК препятствует транскрипции, белок Mfd смещает РНК-полимеразу и рекрутирует комплекс Uvr (см. рис. 24.19 в разд. 24.12 «Связь между транскрипцией и репарацией).

20.4 Способы эксцизионной репарации в клетках млекопитающих

Ключевые понятия:

- **У млекопитающих запуск эксцизионной репарации производится прямым удалением поврежденного основания ДНК.**
- **Удаление основания провоцирует удаление и замену нуклеотидного отрезка.**
- **Какой из двух способов эксцизионной репарации будет активирован, определяется природой реакции удаления основания.**
- **Путь polδ/ε приводит к замене полинуклеотидного отрезка; путь polβ приводит к замене одного нуклеотида.**

Общий принцип эксцизионной репарации в клетках млекопитающих такой же, как у бактерий. Но начинается процесс обычно по-другому — с удаления одного поврежденного основания. Такое начало служит пусковым механизмом активации ферментов, которые вырезают и заменяют отрезок ДНК, содержащий поврежденный сайт.

Ферменты, удаляющие из ДНК азотистые основания, называют гликозилазами и лиазами. На рис. 20.10 показано, как гликозилаза разрезает связь между поврежденным ибо неспаренным основанием и дезоксирибозой. На рис. 20.11 показано, что некоторые гликозилазы являются также лиазами, которые уводят реакцию на шаг вперед, используя аминогруппу (NH_2) для атаки дезоксирибозного кольца. За этим обычно следует реакция внесения надреза в полинуклеотидную цепочку.

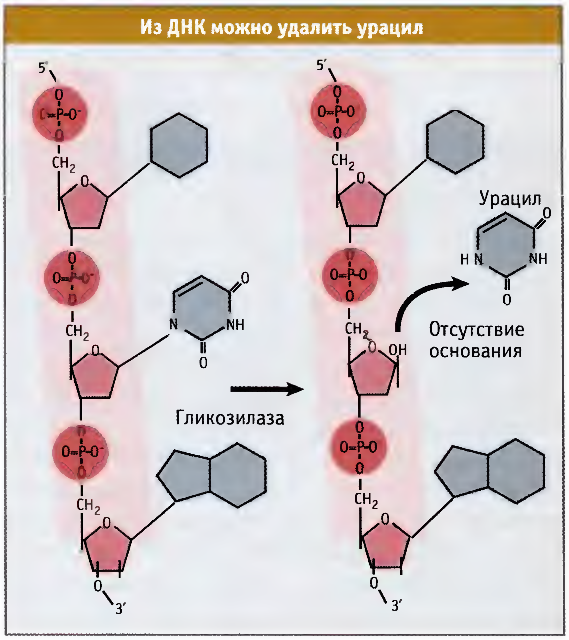


РИС. 20.10 Гликозилаза, разрезая связь азотистого основания с дезоксирибозой, удаляет его из ДНК

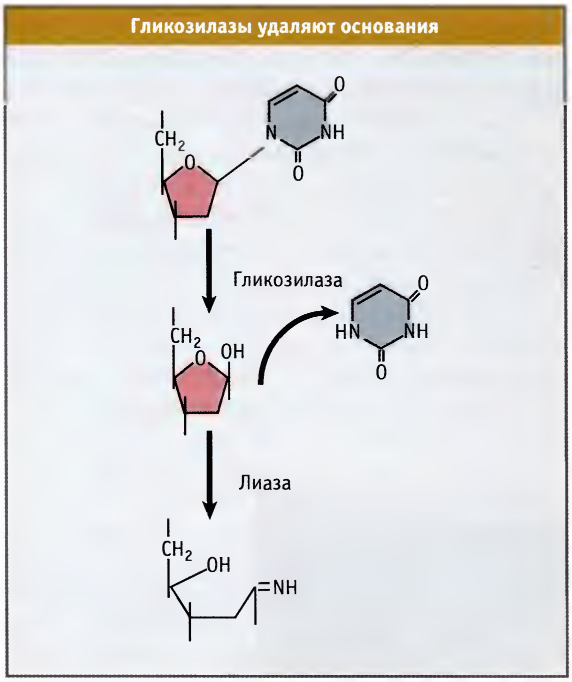


РИС. 20.11 Гликозилаза, расходуя молекулу воды, гидролизует связь между дезоксирибозой и азотистым основанием. Лиаза уводит реакцию дальше: открывает пятичленник сахара при помощи аминогруппы

РИС. 20.12 показывает, что способ репарации в точности зависит от того, какой именно из ферментов, гликозилаза или лиаза, удалил поврежденное основание.

Активность гликозилазы «подхватывает» эндонуклеаза APE1, которая разрезает полинуклеотидную цепочку с 5′-стороны. Это, в свою очередь, привлекает репликативный комплекс, состоящий из ДНК-полимеразы δ/ε и вспомогательных компонентов, который осуществляет ник-трансляцию на расстояние от двух до десяти нуклеотидов. Вытесненный при этом материал удаляется эндонуклеазой FEN1. Фермент лигаза-1 заклеивает цепь. Этот механизм называют «способом длинной заплатки».

Если в инициирующем удалении азотистого основания задействована лиаза, то эндонуклеаза APE1 объединяется с ДНК-полимеразой β для замены единственного нуклеотида. Оставшийся разрыв заклеивает XRCC1/лигаза-3. Такой механизм называют «способом короткой заплатки».

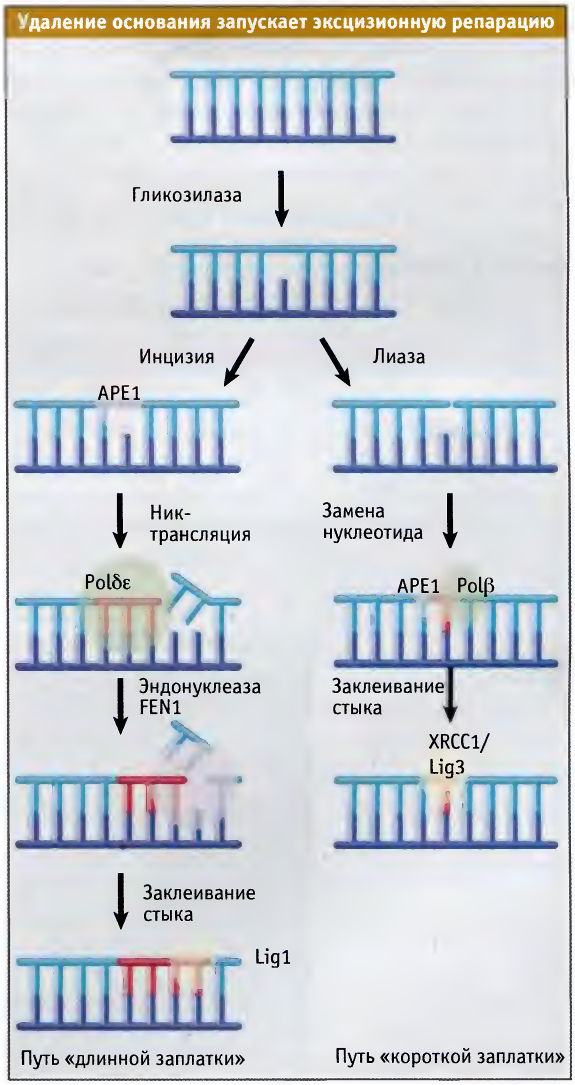


РИС. 20.12 Удаление азотистого основания гликозилазой или лиазой направляет эксцизионную репарацию по одному из двух путей

20.5 Метилазы и гликозилазы вынимают основания из ДНК

Ключевые понятия:

- Гликозилазы узнают и напрямую удаляют из ДНК урацил и алкилированные основания.
- Исправление пиримидиновых димеров производится разрывом ковалентной связи между ними.
- Метилазы добавляют к цитозину метильную группу.
- Все эти ферменты действуют с выведением азотистого основания за пределы двойной спирали, там они удаляют либо модифицируют его и возвращают в спираль.

Несколько ферментов, удаляющих или модифицирующих отдельные основания ДНК, делают это посредством замечательной реакции, при которой азотистое основание «выскакивает» за пределы двойной спирали. Этот тип взаимодействий был впервые обнаружен у метилтрансфераз — ферментов, добавляющих метильную группу к цитозину ДНК. Метилаза вынимает цитозин за пределы двойной спирали. На РИС. 20.13 показано, как он входит в полость фермента, где подвергается модификации. Затем основание возвращается в свое нормальное положение внутри спирали. Все это происходит в отсутствие дополнительного источника энергии.

Одну из самых распространенных реакций, при которых основание удаляется из ДНК, катализирует урацил-ДНК-гликозилаза. Обычно урацил появляется в ДНК в результате (спонтанного) дезаминирования цитозина. Гликозилаза его узнает и удаляет. Реакция похожа на метилазную: урацил выскакивает из спирали в активный сайт гликозилазы. Похожим образом работают и другие гликозилазы и лиазы, в том числе узнающие поврежденные основания ферменты млекопитающих.

Механизм удаления алкилированных оснований (чаще всего тех, к которым была добавлена метильная группа) имеет сходный принцип. У человека один и тот же фермент, алкиладенин-ДНК-гликозилаза (ААГ), узнает и удаляет целый спектр алкилированных субстратов, в том числе 3-метиладенин, 7-метилгуанин и гипоксантин.

В отличие от них 1-метиладенин корректируется ферментом, который использует окислительный механизм. Метильная группа (CH_3) окисляется до CH_2OH, и отделение формальдегида (НСНО) возвращает исходный аденин. У *E. coli* фермент кодируется геном *alkB*; его гомологи (в том числе три гена человека) широко распространены у живых организмов. Интересным оказалось открытие, что бактериальный фермент, а также один из ферментов человека, могут «лечить» такую же поломку азотистого основания в РНК (у человека рибосомальная РНК, возможно, преобладает как субстрат для этого фермента). Это первый обнаруженный вариант репарации РНК.

Другой работающий с выносом основания фермент, фотолиаза, диссоциирует сшивки в пиримидиновых димерах (см. рис. 20.5). Пиримидиновый димер выскакивает в полость внутри фермента. Неподалеку от этой полости расположен активный сайт, который содержит донор электронов для разрыва связей. Реакция идет с поглощением энергии в форме световых волн видимого диапазона.

Общим для этих ферментов является бросок основания-мишени внутрь белковой структуры фермента. Вариацию на ту же тему использует Т4 эндонуклеаза V, в настоящее время переименованная по способу действия в Т4 пиримидин-димер-гликозилазу (Т4 ПДГ). Фермент выносит за пределы двойной спирали аденин, комплементарный тимину на 5′-стороне димера. Мишень катализа — другая половина димера — остается в дуплексе; таким образом, этот фермент использует вынос основания как непрямой механизм доступа к субстрату.

За удалением в ДНК азотистого основания следует эксцизия фосфодиэфирного остова эндонуклеазой, синтез ДНК ДНК-полимеразой для заполнения образовавшейся бреши и заклеивание стыков лигазой для восстановления целостности полинуклеотидной цепочки. Это было описано в предыдущем разделе (см. разд. 20.4 «Способы эксцизионной репарации в клетках млекопитающих»).



РИС. 20.13 Чтобы модифицировать цитозин, метилаза вынимает его из двойной спирали. Любезно предоставлено Sanjay Kumar, New England Biolabs

20.6 Ошибки репаративного синтеза и фенотипы мутаторов

Ключевые понятия:

- Во время репликации поврежденная и нерепарированная ДНК приводит к застреванию ДНК-полимеразы III.
- Цепь, комплементарную поврежденной, могут синтезировать ДНК-полимераза V (кодируемая *umuCD*) или ДНК-полимераза IV (кодируемая *dinB*).
- Последовательность ДНК, синтезированной репаративной ДНК-полимеразой, часто содержит ошибки.
- Гены-мутаторы кодируют белки, от которых зависит точность репликации; мутации мутаторов повышают уровень спонтанного мутагенеза.

В связи с существованием систем репарации, занимающихся синтезом ДНК, возникает вопрос: сравним ли контроль качества нового материала ДНК в таких системах с контролем качества при репликации? Насколько известно, большинство систем, в том числе эксцизионная репарация под контролем *uvr*, по частоте ошибок практически не отличаются от репликации. Однако определенные обстоятельства могут вызвать у *E. coli* другой тип репаративного синтеза ДНК — **«склонный ошибаться»**.

Предрасположенность к внесению в последовательность ДНК ошибок в процессе восстановления повреждений впервые обнаружила себя, когда оказалось, что репарация ДНК фага λ сопровождается индукцией мутаций в случае внесения фага в клетки, предварительно облученные УФ. По всей видимости, УФ-облучение хозяина активировало какие-то клеточные функции, порождающие мутации. Позднее выяснилось, что мутагенный ответ в этом случае затрагивает и бактериальную ДНК хозяина.

Какова действующая причина «склонности к ошибкам»? Эта причина — ДНК-полимераза, которая вставляет в дочернюю цепь ДНК неверный нуклеотид, порождая мутацию. Так бывает, когда фермент оказывается не в состоянии поставить верный, руководствуясь принципом комплементарности. Активности, задействованные при этом, оказались нарушены у бактерий-мутантов по генам *umuD* и *umuC*. У таких мутантов УФ-облучение не индуцирует мутагенез, это означает, что в клетках дикого типа белки UmuC и UmuD вызывают УФ-индуцированный мутагенез. Оперон *umuDC*, содержащий оба гена, побуждается к экспрессии поврежденностью ДНК. Продукты экспрессии образуют комплекс $UmuD'_2C$, который состоит из двух субъединиц укороченного белка UmuD и одной субъединицы UmuC. UmuD обрезается при участии белка RecA активируемого поврежденностью ДНК (см. разд. 20.10 «RecA запускает систему SOS").

Комплекс $UmuD'_2C$ обладает ДНК-полимеразной активностью. Он называется ДНК-полимеразой V и отвечает за синтез новой ДНК на замену УФ-поврежденной. Это единственный фермент *E. coli*, способный прочитывать матрицу, минуя пиримидиновые димеры (классическое УФ-последствие) и другие громоздкие аномалии. Полимеразная активность «склонна ошибаться», и процент неверных нуклеотидов в последовательности дочерней цепи высок. Мутации по *umuC* или *umuD*, инактивирующие фермент, летальны в условиях УФ-облучения. Некоторые плазмиды являются носителями генов *mucA* и *mucB*, гомологичных *umuD* и *umuC*. Ввод таких плазмид в бактерию усиливает и сопротивляемость гибели от УФ, и подверженность мутагенезу.

Как альтернативная ДНК-полимераза получает доступ к ДНК? Когда репликаза (ДНК-полимераза III) встречает на своем пути преграду, такую как тиминовый димер, она застревает. Тогда она, смещаясь за пределы репликативной вилки, уступает место ДНК-полимеразе V. На самом деле, ДНК-полимераза V использует отчасти те же самые вспомогательные белки, что и репликаза. Так же поступает ДНК-полимераза IV, продукт *dinB*, тоже работающая с поврежденной ДНК. ДНК-полимеразы IV и V, вместе с рядом эукариотических ферментов, относятся к большому семейству ДНК-полимераз, участвующих в репарации (см. разд. 20.11 «Системы репарации эукариотических клеток консервативны»).

20.7 Контроль направления репарации несовершенных пар оснований

Ключевые понятия:

- Гены *mut* кодируют систему репарации некомплементарностей по отдельным парам оснований.
- Выбор, в какой из цепей произвести замену нуклеотида для восстановления комплементарности, неслучаен.
- В полуметилированном $\frac{\text{GATC}}{\text{CTAG}}$ обычно заменяется неметилированная цепь.
- Эта система репарации используется для удаления ошибок в новосинтезированной цепи ДНК. В случае неспаренных G-T или C-T предпочтительно удаляется T.

Гены, продукты которых контролируют точность синтеза ДНК в ходе репликации либо репарации, промаркированы мутациями с **мутаторным** фенотипом. Мутант-мутатор характеризуется высокой частотой спонтанного возникновения мутаций. Ген, исходно отождествленный с мутаторным фенотипом, обозначают как *mut*. При этом его продукт впоследствии часто оказывается уже известным компонентом аппарата репликации или репарации.

Типы ферментативных активностей, обусловленных генами *mut*, распадаются на две группы:

- Группа, представленная большинством, состоит из компонентов систем репарации несовершенных пар оснований. Неспособность удалить поврежденное или неспаренное основание до репликации позволяет ему индуцировать мутацию. К ферментам этой

группы относят метилазу *dam*, идентифицирующую мишень для репарации, а также ферменты, прямо или косвенно задействованные в удалении конкретных типов повреждений (*mutH*, *mutS*, *mutL*, *mutY*).
- Меньшая группа, типичный представитель которой — *dnaQ*-кодируемая субъединица ДНК-полимеразы III, имеет отношение к аккуратности при синтезе новой ДНК.

Удаление из ДНК очага искажения ее структуры приводит к восстановлению нуклеотидной последовательности дикого типа. В большинстве случаев деформация ДНК связана с возникновением основания, которое противоестественно для ДНК и которое в связи с этим опознается и устраняется системой репарации.

Проблемы возникают, когда мишенью для репарации служит некомплементарная пара (нормальных) азотистых оснований, образовавшаяся вследствие ошибки. У репарационной системы нет никаких внутренних резонов присвоить одному из них дикий тип, признав другое мутантным! Все, что замечает система, — это два неверно спаренных основания, каждое из которых может представлять собой мишень для эксцизионной репарации.

Если вырезается мутировавшее основание, то восстанавливается последовательность дикого типа. Если же вырезанным случается быть исходному (дикого типа) основанию, то новая (мутантная) последовательность оказывается зафиксированной. Однако чаще всего, в плане выбора основания из некомплементарной пары, направление эксцизионной репарации не случайно, а смещено по вероятности в пользу восстановления последовательности дикого типа.

Например, для таких случаев, как дезаминирование 5-метилцитозина до тимина, существует специальная система восстановления правильной последовательности (см. также разд. 1.15 «Большинство «горячих точек» мутагенеза появляется в результате модификации оснований»). Дезаминирование порождает пару G-T, и система, разрушающая такие пары, предрасположена переделывать их в G-C (а не в A-T). Компоненты этой системы, белки MutL и MutS, избирательно удаляют тимин в обнаруженных некомплементарностях G-T и C-T.

Система *mutT,M,Y* борется с последствиями окислительных перегрузок. В этом случае самый типичный вариант химической модификации азотистых оснований — окисление гуанина до 8-оксогуанина. На РИС. 20.14 изображена трехуровневая система репарации. На первом уровне MutT гидролизует поврежденный предшественник 8-оксо-дГТФ, предотвращая встраивание его в ДНК. С окислением гуанина ДНК, партнер которого цитозин, разбирается MutM, как правило, удаляя из пары G-C «больной» 8-оксогуанин. И все-таки может случиться, что окисленный гуанин «переживет» репликацию в составе матрицы ДНК и породит в дочернем дуплексе ошибочную пару с аденином. Этой парой, A-G, займется белок MutY; он исключит из нее аденин. Белки MutM и MutY — гликозилазы, удаляющие азотистое основание прямо из

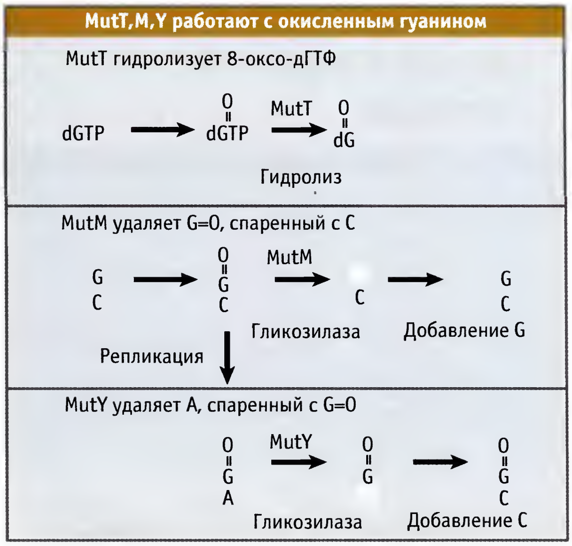


РИС. 20.14 Приоритеты, соблюдаемые при замене оснований в парах, содержащих окисленный гуанин, уменьшают вероятность закрепления мутаций

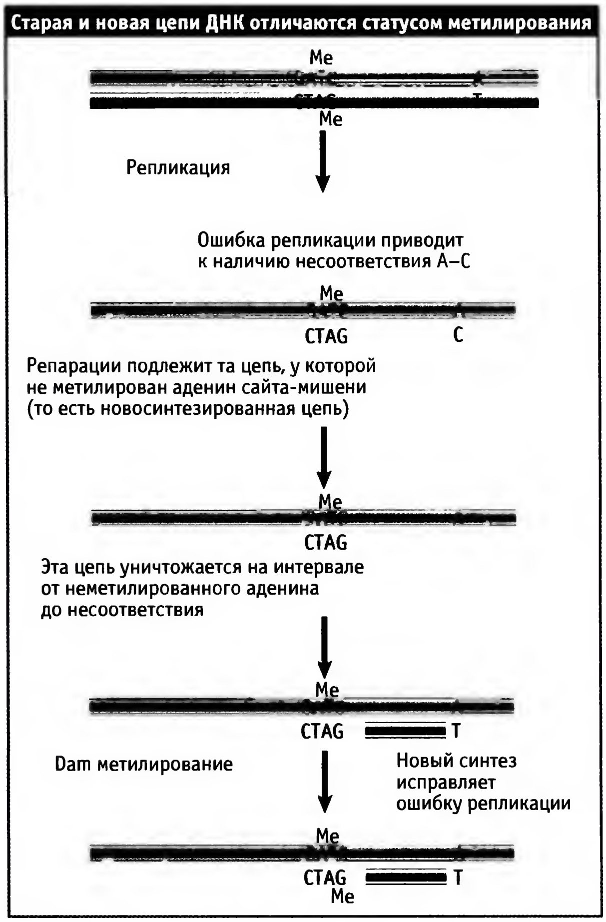


РИС. 20.15 Между раундами репликации наполовину метилированные палиндромы GATC служат мишенями для метилазы Dam. В период времени, проходящий между репликацией и добавочным метилированием, именно неметилированная цепь ДНК является объектом репарации несовершенных пар

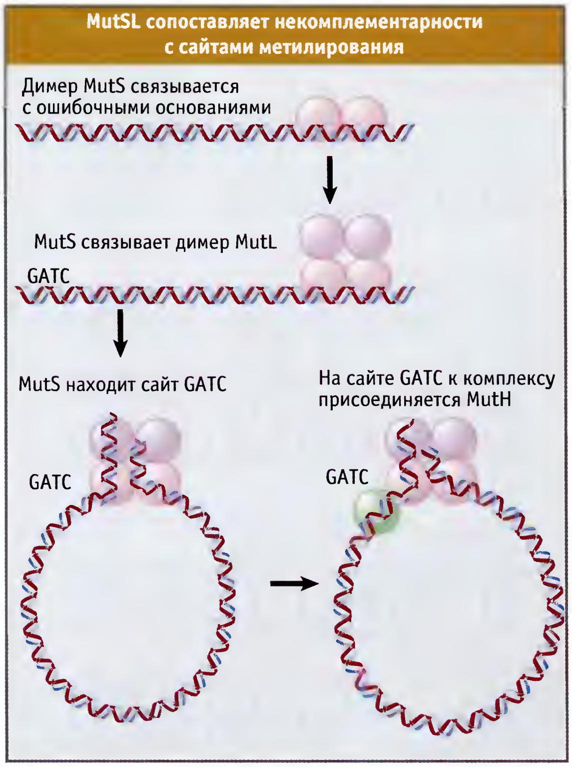


РИС. 20.16 MutS узнает некомплементарность и привлекает к ней ближайший сайт GATC, с которым также связывается. MutH режет неметилированную цепь GATC. После этого эндонуклеазы разрушают неметилированную цепь на интервале от GATC до ошибочного основания

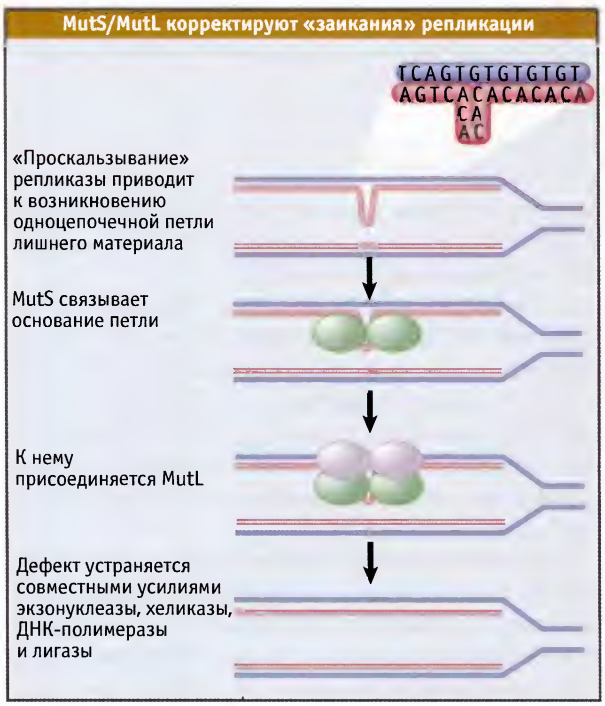


РИС. 20.17 Система MutS/MutL инициирует репарацию несоответствий, возникающих из-за «проскальзывания» репликазы по матрице из частых повторов

ДНК. После удаления (апуринизации) остается дырка, которую узнает эндонуклеаза, срабатывающая на запуск эксцизионной репарации.

Для обработки некомплементарностей, порожденных репликацией, нужен способ отличать дочернюю цепь ДНК от матричной. При репликации ДНК *E. coli* процесс метилирования дочерней цепи отстает от процесса ее синтеза на метилированной матрице. Ген *dam* кодирует метилазу, чья мишень — аденины в обеих цепях палиндрома $\frac{\text{GATC}}{\text{CTAG}}$ (см. рис. 15.7). В промежуток времени, пока дочерняя цепь ожидает метилирования, цепи нового дуплекса различимы по происхождению. Наполовину метилированное состояние сайта GATC используется, например, для различения реплицированных и нереплицированных ориджинов. Такие же наполовину метилированные сайты используют репликационнозависимые системы репарации.

На РИС. 20.15 показано, как при репарации ДНК, содержащей некомплементарные основания, эксцизия «предпочитает» неметилированную цепь. Эффект довольно дальнодействующий: предпочтение остается в силе по >1 тпн окрестности сайта GATC. В результате новосинтезированная цепь приводится в соответствие матрице, а не наоборот.

Мутанты *dam*$^-$ *E. coli* демонстрируют повышенный фон спонтанного мутагенеза. Значит, участие продукта *dam* эффективно помогает некой системе репарации снизить число мутаций, причиняемых ошибками репликации. Система эта состоит из нескольких белков, кодируемых генами *mut*. Белок MutS связывается с некомплементарностью, где к нему присоединяется MutL; MutS имеет два сайта связывания ДНК, это иллюстрирует РИС. 20.16. Один из них специфично распознает некомплементарности, а второй — неспецифичен относительно последовательности или структуры — он используется белком для перемещения вдоль ДНК до встречи с GATC. При перемещении используется гидролиз АТФ. Перемещаясь, MutS продолжает держаться за обнаруженную им некомплементарность, и в результате между некомплементарностью и GATC образуется петля ДНК.

Узнавание последовательности GATC заставляет эндонуклеазу MutH связаться с MutSL. MutH режет неметилированную цепь. Затем эта цепь подвергается эксцизии от GATC до некомплементарности. Эксцизия может идти в направлении 5′–3′ (с участием RecJ или экзонуклеазы VII) или в направлении 3′–5′ (с участием экзонуклеазы I), при поддержке хеликазы UvrD. Новую цепь синтезирует ДНК-полимераза III.

Система репарации *MSH* у *Saccharomyces cerevisiae* гомологична системе *mut E. coli*. Белок Msh2 служит опорной структурой аппарата опознания некомплементарностей, а Msh3 и Msh6 служат факторами специфичности. Комплекс Msh2-Msh3 связывает неспаренные петли размером до четырех нуклеотидов; комплекс Msh2-Msh6 связывается с однонуклеотидными некомплементарностями, а также со вставками

или делециями. Вслед за этим собственно репарация требует участия других белков.

Гомологи системы MutSL также обнаружены в клетках высших эукариот. В дополнение к репарации однонуклеотидных некомплементарностей, они ответственны за устранение дефектов, возникающих вследствие «проскальзывания» репликации. В районе, где имеет место многократный повтор очень короткой последовательности (микросателлит), пересовмещение дочерней цепи и матрицы может повлечь за собой «заикание», при котором ДНК-полимераза проскальзывает назад и синтезирует лишние повторы. Эти повторы в составе дочерней цепи выпирают из двойной спирали в виде одноцепочечных петель (см. рис. 6.28). Их репарацией занимаются системы, гомологичные MutSL; это изображено на РИС. 20.17.

О важности системы MutSL для репарации некомплементарностей говорит высокая доля онкологических заболеваний при нарушении ее функции. Потеря клеткой этой системы ведет к повышенной частоте возникновения мутаций.

20.8 Системы рекомбинационной репарации *E. coli*

Ключевые понятия:

- Гены rec *E. coli* кодируют компоненты системы заимствования информации из другого дуплекса.
- Система заимствования информации из другого дуплекса приходит в действие, когда репликация оставляет в новой цепи ДНК брешь напротив поврежденного участка матрицы.
- Чтобы восполнить пробел, используется одиночная цепь другого дуплекса.
- Затем поврежденная последовательность удаляется и синтезируется заново.

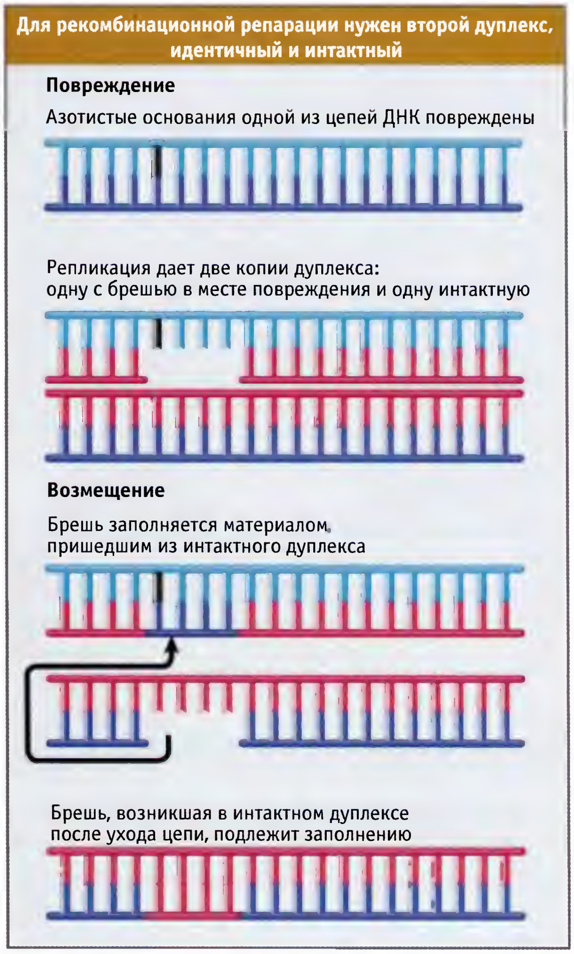


РИС. 20.18 Работа системы принципиального возмещения *E. coli* основана на заимствовании верной генетической информации, полностью утраченной в одной геномной копии, из другой, интактной, копии

Системы рекомбинационной репарации задействуют активности ферментов, участвующих и в генетической рекомбинации. Иногда их называют системами пострепликативной репарации, т. к. они приступают к работе после репликации. Такие системы эффективно борются с дефектами дочерней цепи, порожденными репликацией поврежденной матрицы. Пример изображен на РИС. 20.18. Возможно, главная задача рекомбинационной репарации состоит в перезапуске застрявших репликативных вилок (см. разд. 18.17 «Праймосома перезапускает репликацию»).

Представим себе, как пиримидиновый димер, находясь на одной из сторон двойной спирали ДНК, коверкает ее структуру. При репликации димер не дает оккупированному им сайту служить матрицей. Репликация бывает вынуждена проскочить его.

Вероятно, ДНК-полимераза подходит вплотную (во всяком случае, близко) к пиримидиновому димеру. После этого полимераза прекращает синтезировать дочернюю цепь. Проскочив на некоторую дистанцию вперед, фермент стартует вновь, оставляя в новосинтезированной цепи довольно крупную брешь.

В результате после репликации дочерние дуплексы отличаются. В одном из них матричная цепь содержит поврежденное звено напротив крупной бреши в дочерней цепи. С другим дуплексом все в порядке: интактная матричная цепь была беспрепятственно скопирована в виде обращенной к ней комплементарной дочерней цепи. Наличие нормальной цепи использует система заимствования информации.

Брешь напротив поврежденного сайта в первом дуплексе заполняется гомологичным участком цепи ДНК, заимствованным у нормального дуплекса. После такой **одноцепочечной перестановки** у дуплекса-реципиента поврежденная матричная цепь оказывается обращенной к интактной цепи дикого типа. А дуплексу-донору, в пару к нормальной родительской цепи и взамен полноценной дочерней, достается цепь, содержащая брешь, которая может быть заполнена

обычным способом репарационного синтеза. Таким образом, общий ущерб минимизируется до исходной деформации (хотя те же приемы рекомбинационной репарации придется воспроизводить после каждого цикла репарации, если — и до тех пор, пока — пиримидиновый димер (или другой дефект), вызывающий искажение структуры ДНК, не будет устранен системой эксцизионной репарации).

Доминирующий путь рекомбинационной репарации *E. coli* обусловлен генами *rec* (см. рис. 19.13–19.15). В клетках *E. coli*, испытывающих недостаточность эксцизионной репарации, мутация гена *recA*, перекрывает практически все остаточные репарационные и ДНК-восстанавливающие мощности. В клетках с двойной мутацией $uvr^{-}recA^{-}$ попытки реплицировать ДНК приводят к образованию фрагментов, чьи размеры соответствуют ожидаемому промежутку между тиминовыми димерами. Этот результат позиционирует тиминовые димеры как летальную преграду для репликации в отсутствие функции RecA. Это объясняет, почему двойные мутанты не выдерживают более двух димеров в своем геноме (по сравнению со способностью бактерии дикого типа справиться с пятьюдесятью).

Один из путей rec задействует гены *recBC* и изучен хорошо; другой задействует *recF* и пока не так хорошо понят. В условиях *in vivo* они служат разным надобностям. Путь *recBC* применяется для «рестарта» застрявших репликативных вилок (см. разд. 20.9 «Рекомбинация служит важным механизмом исправления ошибок репликации»). Путь *recF* участвует в репарации брешей дочерней цепи, оставшихся после ее синтеза в обход пиримидинового димера.

Пути *recBC* и *recF* по-разному предшествуют действию RecA. Оба пути приводят к ассоциации RecA с одиночной цепью ДНК. Способность RecA к перестановке одиночных цепей позволяет ему «разрулить» этап заимствования информации (рис. 20.18), после чего нуклеаза и полимераза могут довести репарацию до конца.

Путь *recF* сочетает в себе функции трех генов, *recF*, *recO* и *recR*. Их белки образуют два типа комплексов, RecOR и RecOF. Они способствуют образованию филаментов RecA на одноцепочечной ДНК, и одна из этих функций позволяет филаментам образовываться, несмотря на ингибирующее присутствие белка SSB. Полагают, что их действие привязано к брешам; вместе с тем реакция *in vitro* требует свободного 5′-конца.

Обозначения генов-участников репарации и рекомбинации базируются на фенотипах мутантов. Иногда бывает, что мутация, изолированная при одних условиях как маркер локуса, скажем, *uvr*, была изолирована ранее, при другом наборе условий, как маркер-идентификатор локуса *rec*. У этой неопределенности есть глубокий смысл. Пока что нельзя сказать, сколько функций принадлежит какому пути, и как пути взаимодействуют между собой. Пути *uvr* и *rec* не полностью взаимно обособлены, поскольку у мутантов *uvr* эффективность рекомбинационной репарации снижена. В конце концов, следует ожидать обнаружения функциональной сети из нуклеазной, полимеразной и других активностей, которые динамически сочетаются в составах различных систем репарации, перекрывающихся по компоновке — в плане возможностей замены какого-либо фермента аналогом из другого пути.

20.9 Рекомбинация служит важным механизмом исправления ошибок репликации

Ключевые понятия:

- **Репликативная вилка может застрять при встрече в ДНК поврежденного участка или ника.**
- **Репарация застрявшей вилки может включать в себя спаривание новосинтезированных цепей.**
- **Застрявшая вилка может заново стартовать после репарации повреждения, используя для продвижения хеликазу.**
- **Застрявшая вилка представляет собой структуру Холлидея и может быть разведена на дуплексы разрывом двух цепей при участии резольваз.**

Каждая клетка владеет несколькими способами починки поврежденной ДНК. Выбор способа обусловлен типом повреждения, а также обстоятельствами. Пути эксцизионной репарации могут быть использованы в любой момент времени, а вот рекомбинационная репарация доступна к применению только при наличии второго дуплекса с копией поврежденной последовательности, то есть сразу после репликации. Особый случай представляет репликация поврежденной ДНК, потому что репликативная вилка может застрять в дефектном сайте. Пути рекомбинационной репарации могут помочь восстановить репликативную вилку после репарации повреждения или позволить репликативной вилке обойти повреждение.

На рис. 20.19 изображен один из возможных исходов застревания репликативной вилки. «Столкнувшись» с повреждением, вилка прекращает движение. Репликативный аппарат подвергается разборке, по крайней мере частичной. Это делает возможной миграцию ветви со сдвигом вилки назад, причем новосинтезированные дочерние цепи встают в пару, образуя дуплекс. После репарации повреждения хеликаза «прокатывает» ветвь вперед, структура ДНК вилки восстанавливается, репликативный аппарат монтируется заново, и репликация перезапускается (см. разд. 18.17 «Праймосома перезапускает репликацию»). Для перезапуска требуется ДНК-полимераза II, потом ее заменяет ДНК-полимераза III.

Возобновление работы застрявшей репликативной вилки требует участия ферментов репарации. У *E. coli* важную роль в этой реакции играют системы RecA и RecBC (возможно, это их главная — для бактерии — роль). Один из возможных способов — стабилизация белком RecA одноцепочечной формы ДНК посредством связывания ее в застрявшей репликативной вилке и, возможно, выступление при этом RecA в качестве сенсора, детектирующего застревание. RecBC участвует в эксцизионной репарации повреждения. После ремонта ДНК репликация может быть возобновлена.

Другой путь проходит через использование рекомбинационной репарации, возможно, с реакцией

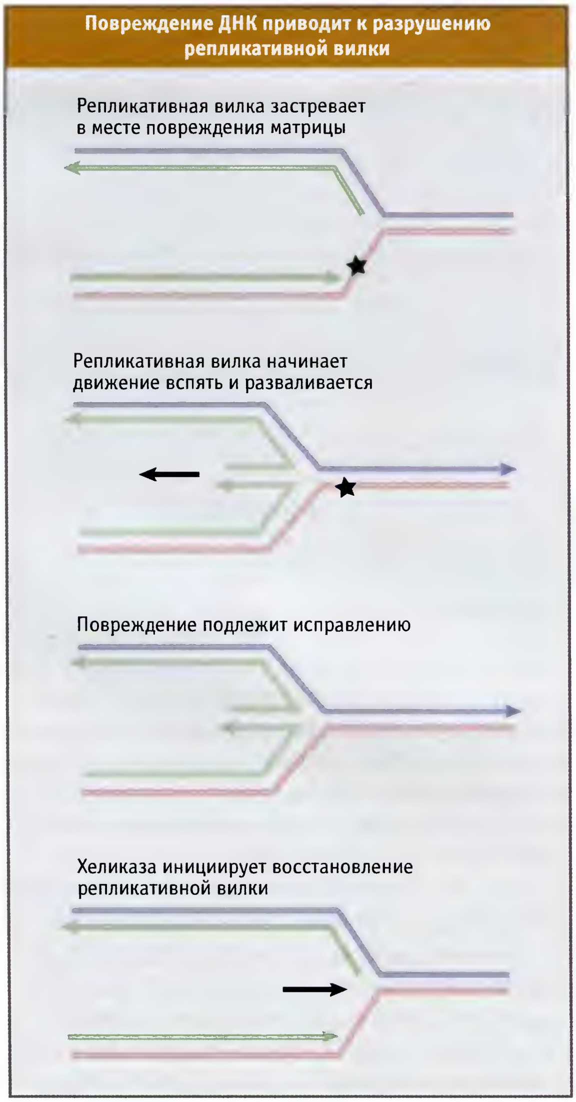


РИС. 20.19 Дойдя до места повреждения матрицы, репликативная вилка застревает. При обратном ходе вилки две комплементарные дочерние цепи образуют дуплекс. После того как повреждение исправлено, хеликаза катализирует миграцию ветви вперед: репликативныая вилка восстанавливается, репликация продолжается. Стрелками обозначены 3′-концы



РИС. 20.20 Застрявшую репликативную вилку можно считать эквивалентом структуры Холлидея, и резольвазы могут осуществлять ее разрезание. Результат разрезания зависит от природы повреждения, остановившего ход репликативной вилки. Не является ли оно разрывом само по себе? Если нет, резольвазы внесут в ДНК один двухцепочечный разрыв. Если да, то в месте повреждения появится второй двухцепочечный разрыв. Стрелками обозначены 3′-концы

перестановки цепей ДНК белком RecA. На РИС. 20.20 структура застрявшей репликативной вилки идентична структуре Холлидея, представляющей собой продукт рекомбинации между двумя дуплексами ДНК. Это делает ее мишенью для резольваз. Разрезание резольвазой любой пары одинаковых цепей порождает двухцепочечный разрыв. Если вдобавок само повреждение содержало ник, то получается, что действие резольвазы породило два двухцепочечных разрыва.



РИС. 20.21 Если репликативная вилка застряла, аппарат репарационной рекомбинации может закрыть поврежденный сайт матрицы интактной цепью. Это способствует возобновлению репликации

Застрявшие репликативные вилки могут быть спасены рекомбинационной репарацией. Как в точности развиваются события, неизвестно; один из возможных сценариев дан на РИС. 20.21. Принцип же состоит в том, что событие рекомбинации охватывает оба фланга окрестности ущербного сайта, позволяя интактной одноцепочке образовать пару с поврежденной цепью. Это позволяет репликативной вилке реконструироваться и продолжить репликацию, удачно миновав поврежденное место.

20.10 RecA запускает SOS-систему

Ключевые понятия:

- Повреждение ДНК приводит к тому, что белок RecA запускает SOS-ответ, обусловленный ферментами-продуктами многочисленных генов.
- RecA активирует саморасщепление LexA.
- LexA репрессирует систему генов SOS; его саморасщепление приводит к активации SOS-системы

Непосредственное участие в рекомбинационной репарации — лишь одна из многих активностей белка RecA. Этот необычайный белок имеет и другую, совершенно отдельную, функцию. Эта функция может быть активирована разного рода воздействиями, которые повреждают ДНК или ингибируют репликацию у *E. coli*. В этом случае RecA срабатывает на запуск сложной серии фенотипических изменений под общим названием **SOS-ответ**. Этот ответ включает экспрессию многих генов, в том числе — ферментов репарации. Из-за функциональной двойственности белка RecA бывает трудно однозначно объяснить, чем обусловлен дефицит репарации у мутантов *recA*⁻: неспособностью RecA к перестановке цепей ДНК или же потерей какой-то другой клеточной функции, чья индукция обусловлена сопряженной с RecA протеазной активностью.

Индуцирующее SOS-ответ повреждение ДНК может быть результатом УФ-облучения (наиболее изученный случай) или действия химических агентов («сшивающих», алкилирующих и т. п.). Ингибирование репликации одним из способов — включая тиминовое голодание, действие лекарств, мутации по нескольким генам *dna* — имеет тот же эффект.

Клеточный SOS-ответ, в форме усиления «репаративной мощи», достигается индукцией синтеза компонентов-участников двух путей: эксцизионной репарации длинными заплатками и RecA-обусловленной рекомбинационной репарации. Кроме того, в условиях SOS-ответа подавлено клеточное деление; могут быть индуцированы лизогенные профаги.

Весь SOS-ответ проистекает из перепрофилирования функции RecA вследствие вредоносного воздействия среды. Как именно оно обусловливает внезапное изменение в активности RecA, не очень понятно. Повлечь за собой SOS-ответ может целый спектр вредоносных воздействий, и в последние годы работа фокусируется на выявлении гипотетического активатора RecA среди общих промежуточных продуктов метаболизма ДНК.

Сигнал может представлять собой небольшую молекулу, которая отделяется от ДНК, или же поступать в форме структуры, образующейся в самой ДНК. Активация RecA в опытах *in vitro* требует наличия одноцепочечной ДНК и АТФ. Возможно, активирующим сигналом служит наличие участка одиночной цепи по месту повреждения. Сигнал, в какой бы форме он ни поступал, взаимодействует с RecA очень быстро: SOS-ответ срабатывает уже в первые несколько мин после обработки.

Активация RecA вызывает протеолитическое расщепление продукта гена *lexA*. LexA, небольшой (22 кДа) белок, в необработанных клетках относительно стабилен и работает репрессором многих оперонов. Реакция его расщепления необычна: LexA обладает латентной протеазной активностью, которую активирует RecA. Когда RecA стресс-активирован, он заставляет LexA предпринять автокаталитическое расщепление; это инактивирует LexA как репрессор, тем самым индуцируя все опероны, с которыми тот был связан. Этот путь изображен на РИС. 20.22.

Многие из генов, в стандартных условиях репрессируемые LexA, кодируют ферменты репарации. Некоторые из этих генов активны лишь в обработанных клетках; другие были активны и в покое, но с расщеплением LexA уровень их экспрессии повышается. Например, у гена *uvrB*, компонента системы эксцизионной репарации, есть два промотора: один подчинен контролю LexA, а другой работает независимо. После расщепления LexA ген может быть экспрессирован не только со второго, но и с первого промотора.

LexA репрессирует свои мишени, связываясь с отрезком в 20 пн, называемым **SOS-боксом**, содержащим консенсус из восьми абсолютно консервативных позиций. Обычно SOS-бокс лежит внутри промотора. Локус *lexA*, подлежащий саморегулирующейся репрессии, содержит два смежных SOS-бокса.

Функциональное сцепление RecA с LexA выражается в том, что RecA запускает расщепление LexA, который репрессировал и *recA*, и свой собственный ген. Таким образом, в рамках SOS-ответа идет наработка обоих (RecA и LexA) белков. Результаты не столь противоречивы, как могло бы показаться на первый взгляд.

Интенсификация продукции белка RecA необходима для его использования по прямому назначению в рекомбинационной репарации. При SOS-индукции экспрессия RecA, с базового уровня в 1200 молекул на клетку, может подскочить в 50 раз. Высокая концентрация RecA в обработанных клетках гарантирует расщепление всего LexA. Это, по идее, должно предотвращать восстановление LexA в роли репрессора целевых генов.

Колебательный контур между RecA и LexA способствует быстрому возвращению клетки к норме при отмене травмирующего воздействия среды. При снятии индуцирующего сигнала белок RecA утрачивает способность дестабилизировать LexA. К этому моменту *lexA* пребывает на пике своей экспрессии; без «нападения» RecA белок LexA начинает быстро накапливаться в нерасщепленной форме и репрессировать SOS-гены. Вот почему SOS-ответ так легко обратим.

RecA также запускает расщепление и других клеточных мишеней, иногда с более прямолинейным эффектом. Так, при активации RecA расщепляется белок UmuD; это расщепление активирует UmuD и «склонную ошибаться» систему репарации. По современным представлениям рабочая модель этой реакции подразумевает связывание комплекса $UmuD_2UmuC$ с RecA-филаментом вблизи места повреждения. RecA активирует комплекс, разрезая

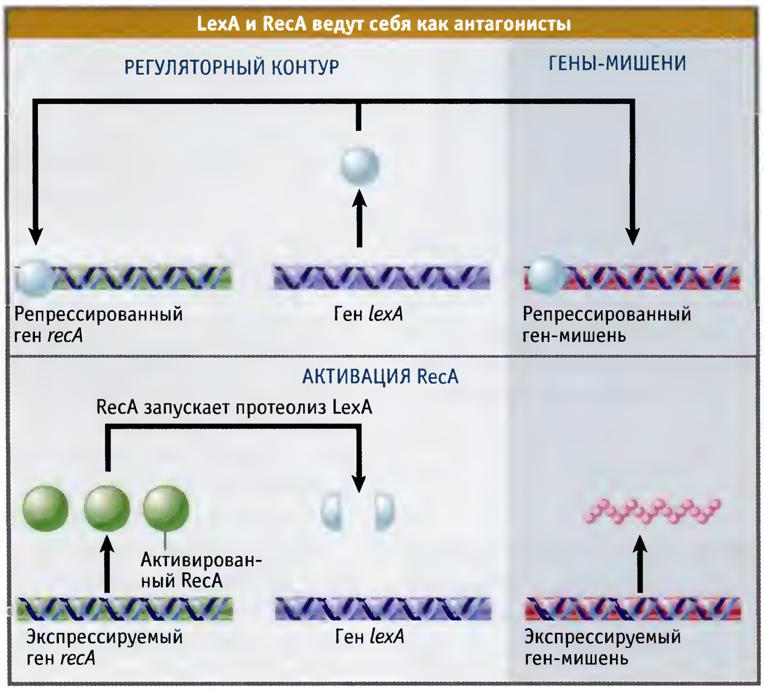


РИС. 20.22 Белок LexA служит репрессором для множества генов, в том числе, обеспечивающих репарационные функции, а также генов *recA* и *lexA*. Активация RecA ведет к автокаталитическому расщеплению LexA, которое, в свою очередь, ведет к индукции экспрессии всех этих генов

UmuD с образованием UmuD', после чего UmuD'$_2$UmuC (ДНК-полимераза V) синтезирует фрагмент ДНК взамен поврежденного материала.

Активация RecA вызывает расщепление репрессоров нескольких профагов (и некоторых других белков-репрессоров). На примере расщепления репрессора лямбда и была в свое время открыта протеазная SOS-активность.

Лизогенный репрессор расщепляется и фаг вступает в литический цикл — таков механизм индукции лямбда ультрафиолетовым облучением. Реакция не относится к клеточному SOS-ответу, но она есть следствие усвоения профагом информации о том, что клетка в беде. Выживанию вируса в таких условиях сообразен выход в литический цикл для генерации потомства. Используя активированный RecA, фаг, по сути, считывает клеточную систему оповещения. Информацию о состоянии клетки-хозяина фаг использует для соблюдения собственных интересов (в целях выживания).

Две функции RecA относительно независимы. При мутации *recA441* SOS-ответ, бывает, срабатывает и в отсутствие стресс-обработки благодаря спонтанной активации RecA. Другие мутации лишают RecA свойства активироваться. Ни одна из перечисленных мутаций не нарушает обращения RecA с ДНК. И если бы вдруг нашлась мутация RecA, которая, наоборот, инактивирует его рекомбинационную функцию, не затрагивая способности индуцировать SOS-ответ, такая мутация весьма пригодилась бы в расследовании прямых и косвенных влияний RecA на репарацию.

20.11 Системы репарации эукариотических клеток консервативны

Ключевые понятия:

- У дрожжей мутации *RAD*, связанные с фенотипом повышенной чувствительности к радиации, находятся в генах, кодирующих системы репарации.
- Болезнь человека ксеродерма пигментная (КП) обусловлена мутациями в нескольких генах репарации.
- Белковый комплекс, в котором состоят продукты КП и фактор транскрипции $TF_{II}H$, выполняет эксцизионную репарацию в клетках человека.
- Транскрипционно-активные гены пользуются приоритетом при репарации.

Разновидности репарационных функций, идентифицированные у *E. coli*, являются общими для широкого круга живых организмов. Лучше других эукариотических систем изучены дрожжи, в которых белок Rad51 функционирует аналогично бактериальному RecA. У дрожжей основное предназначение белка-переносчика цепи — участие в гомологичной рекомбинации. Многие системы репарации дрожжей имеют прямых аналогов в клетках высших эукариот, в связи с неполадками они могут быть причиной заболевания человека.

Гены *RAD* причастны к репарации; благодаря чувствительности к радиации они подробно охарактеризованы у дрожжей с позиций генетики. У *S. cerevisiae* эти гены разбиты на три группы: *RAD3* (гены, задействованные в эксцизионной репарации), *RAD6* (необходимы для пострепликативной репарации), *RAD52* (кодируют составляющие механизмов, аналогичных рекомбинации).

Группа *RAD52* подразделяется на две подгруппы на основании различий мутантных фенотипов. У мутантов одной подгруппы нарушена гомологичная рекомбинация, что выражается в спаде митотической рекомбинации у мутантов по *RAD50*, *RAD51*, *RAD54*, *RAD55* и *RAD57*. Эти белки Rad образуют мультипротеиновый комплекс по месту двухцепочечного разрыва. После того, как экзонуклеаза подействует на свободные концы и породит одноцепочечные «хвосты» по краям пробела, Rad51 инициирует процесс, связывая одноцепочечную ДНК и образуя с ней нуклеопротеиновый филамент. Затем с филаментом последовательно связываются Rad52, Rad55 и Rad54. У мутантов по *RAD59*, *MRE11* и *XRS2* — второй подгруппы *RAD52*-генов — частота рекомбинации, наоборот, повышена. Они не испытывают трудностей с гомологичной рекомбинацией, но дефицитны по негомологичному соединению концов ДНК.

Надсемейство ДНК-полимераз, участвующих в синтезе ДНК взамен поврежденного материала, образуют гены *dinB* и *umuCD* (кодирующие ДНК-полимеразы IV и V) *E. coli*, *RAD30* (кодирующий ДНК-полимеразу η) *S. cerevisiae*, *XPV H. sapiens*. Иногда это надсемейство называют транслесионными ДНК-полимеразами. Эукариотические ферменты, в отличие от бактериальных, прочитывают тиминовые димеры без ошибок, аккуратно ставя два аденина против Т-Т димера. При репликации других типов повреждений они все же «склонны ошибаться».

Интересной особенностью репарации, лучше всего описанной у дрожжей, является ее связь с транскрипцией. Репарация транскрипционно-активных генов приоритетна. Это особенно касается транскрибируемой цепи и направлено на устранение помех для транскрипции. Действующей причиной этого служит, по-видимому, механическая связь между репарационной машиной и РНК-полимеразой. Белок Rad3 — хеликаза, необходимая на этапе внесения надсечек — является компонентом фактора транскрипции, ассоциированного с РНК-полимеразой (см. разд. 24.12 «Связь между транскрипцией и репарацией»).

Клетки млекопитающих гетерогенны по количеству ДНК, синтезируемому в местах разрушений, чтобы закрыть повреждение. Однако каждая в отдельности «заплатка» — довольно короткая, меньше 10 оснований в длину.

О важности систем репарации млекопитающих говорит ряд наследственных заболеваний человека. Наиболее исследованное из них — ксеродерма пигментная (xeroderma pigmentosum, XP), рецессивно наследуемая гиперчувствительность к солнечному свету, в особен-

ности ультрафиолетовой части спектра. XP приводит к поражению кожных покровов (а иногда и более суровым дефектам).

Заболевание связано с нарушением эксцизионной репарации. Фибробласты пациентов с XP не способны к эксцизии пиримидиновых димеров и других громоздких дефектов собственной ДНК. Мутации, вызывающие XP, распределены по восьми генам с названиями от *XP-A* до *XP-G*, соответственно буквам латинского алфавита. Их гомологи есть среди генов *RAD* у дрожжей, так что этот путь распространен у эукариот. Белковый комплекс, содержащий продукты нескольких генов *XP*, отвечает за эксцизию тиминовых димеров. На РИС. 20.23 отмечена ступенька, которую занимает этот комплекс в пути репарации. Комплекс связывает ДНК по месту повреждения; механизм связывания, вероятно, основан на кооперации между несколькими компонентами комплекса. Цепи ДНК расплетаются на участке протяженностью ~20 пн, окружающем повреждение. Это — результат хеликазной активности фактора транскрипции $TF_{II}H$, представляющего собой обширный комплекс (в который входят также продукты нескольких генов *XP*) и участвующего в репарации повреждений ДНК (которые во время транскрипции встречает на своем пути РНК-полимераза). На следующем этапе репарации эндонуклеазы, кодируемые генами *XP*, вносят надрезы (надсечки) по обе стороны расплетенного участка. После этого фрагмент одной из цепей, несущий в себе дефект, может быть заменен свежесинтезированным материалом.

В тех случаях, когда с неудаленным тиминовым димером сталкивается репликация, для того, чтобы она продолжилась, требуется активность ДНК-полимеразы η. Эту полимеразу кодирует ген *XPV*. Рак кожи у мутантов по *XPV* обусловлен, по всей видимости, ее отсутствием.

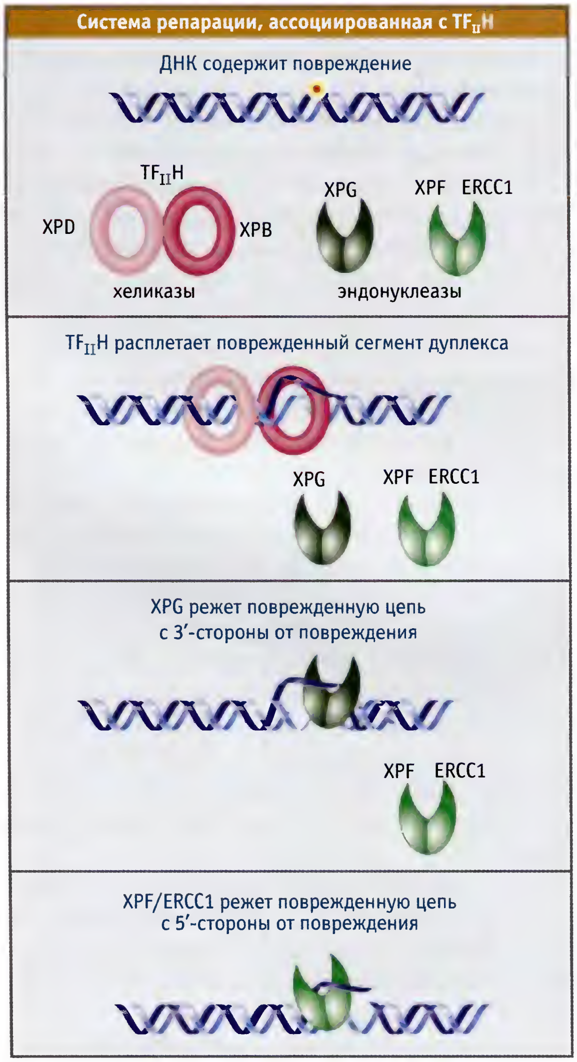


РИС. 20.23 Хеликаза расплетает ДНК в месте повреждения, эндонуклеазы вырезают повреждение, ДНК-полимераза заполняет одноцепочечную брешь

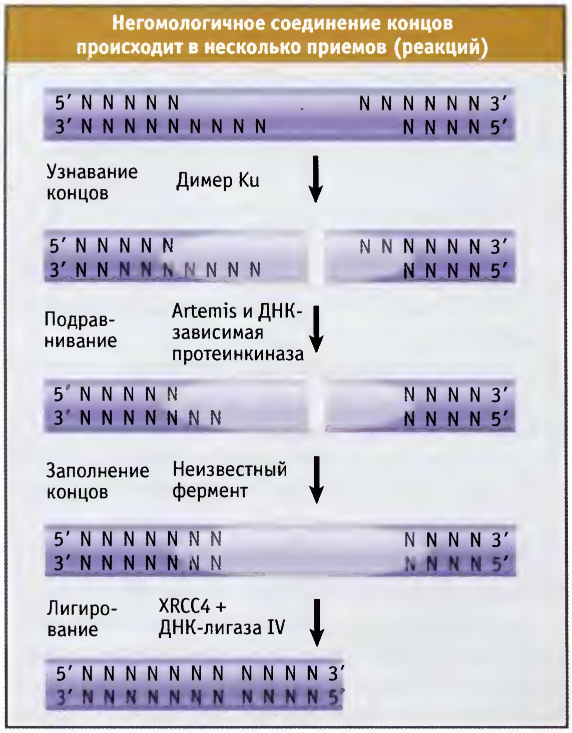
20.12 Неспециализированная система репарации двухцепочечных разрывов

Ключевые понятия:

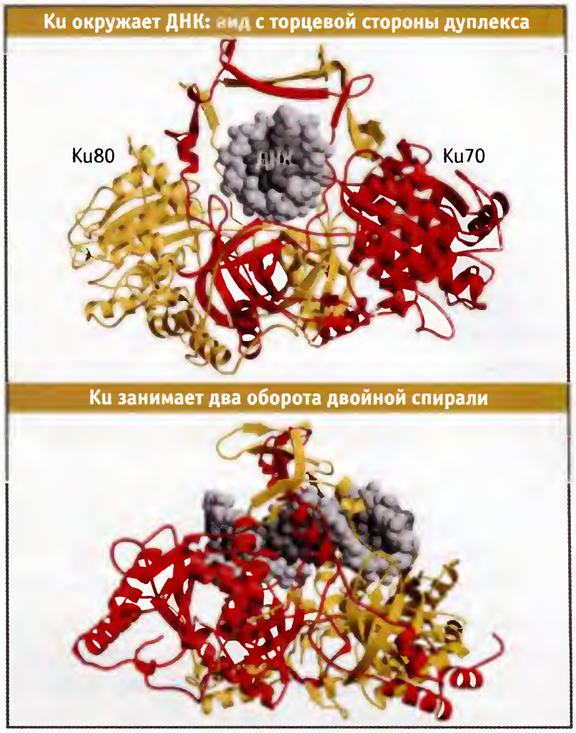
- **Путь NHEJ может склеивать тупые концы дуплексов ДНК.**
- **Мутации, нарушающие путь NHEJ у человека, приводят к заболеваниям.**

Двухцепочечные разрывы случаются в клеточной ДНК при разных обстоятельствах. Они инициируют гомологичную рекомбинацию и фигурируют на промежуточных этапах рекомбинации иммуноглобулиновых генов (см. разд. 23.10 «Белки RAG катализируют разрыв и воссоединение ДНК»). Также они могут представлять собой повреждение ДНК — например, последствия облучения. Главный механизм репарации таких разрывов называется **негомологичным соединением концов (NHEJ)** и заключается в лигазной реакции, которая склеивает вместе тупые концы ДНК.

РИС. 20.24 резюмирует этапы NHEJ. Один и тот же ферментный комплекс берет на себя и «ремонтные работы», и иммунную рекомбинацию. Первый этап заключается в узнавании рваных концов белковым гетеродимером в составе Ku70 и Ku80. Они образуют механическую опору для удержания тупых концов ДНК вместе, предоставляя другим белкам этими концами оперировать. Ключевой участник процесса — ДНК-зависимая протеинкиназа, которую ДНК активирует, побуждая фосфорилировать белковый субстрат. Одна из ее мишеней — белок Artemis, который в активной (фосфорилированной) форме может работать и экзонуклеазой, и эндонуклеазой (срезать шпильки, образовавшиеся при рекомбинации иммуноглобулиновых генов). Что за ДНК-полимераза заполняет все оставшиеся одноцепочечные выступы, неизвестно. Наконец, собственно соединение концов — функция ДНК-лигазы IV, которая



Основные этапы негомологичного соединения концов (NHEJ) включают: узнавание разрывов ДНК, подравнивание и/или заполнение одноцепочечных концевых выступов и лигирование



Гетеродимер Ku70-Ku80 раскинулся на два витка двойной спирали ДНК; в центральной части сайта связывания он окружает ДНК кольцом. Любезно предоставлено Jonathan Goldberg, Memorial Sloan-Kettering Cancer Center

действует в связке с белком XRCC4. Мутации по любому из перечисленных компонентов могут ввести эукариотическую клетку в состояние гиперчувствительности к радиации; некоторые из этих мутаций сопутствуют известным человеческим болезням, обусловленным нарушениями репарации ДНК.

Гетеродимер Ku — это сенсор, детектирующий повреждение ДНК посредством связывания рваных концов. Результат расшифровки его кристаллической структуры — молекулярная модель на . Массив белка прилегает к двойной спирали по одной (нижней) ее стороне, контактируя с ней на протяжении примерно двух витков, однако посередине этого расстояния ДНК опоясана узким мостиком, соединяющим субъединицы димера. Таким образом, гетеродимер похож на корзинку, «ручка» которой может надеться только на свободный конец ДНК.

Ku может свести рваные концы вместе, связав две молекулы ДНК. Судя по способности гетеродимеров к ассоциации между собой, реакция может выглядеть подбно схеме на . Такая модель отводит лигазе место для связывания с ДНК в промежутке, образованном мостиками двух гетеродимеров Ku. После реакции Ku, чтобы освободиться от ДНК, должен изменить свою структуру.

Общая черта заболеваний, вызванных недостатками репарации по типу NHEJ, — нестабильность хромосом в связи с неспособностью репарировать двухцепочечные



РИС. 20.2 Когда два гетеродимера Ku связывают ДНК в комплексе друг с другом, расстояние между их «ручками», обхватывающими двойную спираль, равняется ~12 пн

разрывы. Это проявляется в виде хромосомных аберраций, ассоциированых с высокой частотой мутаций и многократно увеличенным риском развития рака. Основополагающей причиной может быть дисфункция или путей контроля репарации, или ферментов в репаративных комплексах. Симптоматика при этом может быть очень схожей, как у ataxia telangiectasia (AT), обусловленной крушением контрольно-пропускной точки клеточного цикла, так и у синдрома Nijmegan (NBS), обусловленного мутацией фермента репарации.

Консервативность путей репарации среди млекопитающих, дрожжей, бактерий — один из важнейших выводов, сделанных при их изучении. Рецессивно наследуемый синдром Блума обусловлен мутацией в гене хеликазы BLM, гомологичном *recQ E. coli*. В результате мутации повышена частота разрывов и перестановок между сестринскими хроматидами. BLM входит в состав обширного комплекса вместе с другими белками репарации. Один из его партнеров по взаимодействию — hMLH1, белок-участник репарации некомплеметарностей, гомолог бактериального MutL. Дрожжевые гомологи BLM и hMLH1 — Sgs1 и MLH1 — тоже ассоциированы физически, что говорит об эволюционной консервативности пути, к которому они принадлежат.

Синдром Nijmegan — результат мутаций в гене, чей белок называют по-разному: Nibrin, p95 или NBS1. Он — компонент репарационного комплекса Mre11/Rad50. Его участие в репарации двухцепочечных разрывов демонстрирует фокальная концентрация определенных белков в клетках человека при обработке, индуцирующей двухцепочечные разрывы ДНК. После воздействия (облучения) киназа ATMP (кодируемая геном *AT*) фосфорилирует NBS1; это активирует комплекс, локализованный в местах повреждений ДНК. Последующие этапы формирования центров репарации включают запуск контрольно-пропускного механизма, останавливающего клеточный цикл на период репарации ДНК, и привлечение в эти центры остальных белков-участников процесса репарации.

20.13 Резюме

Бактерии содержат системы, поддерживающие целостность их ДНК после повреждений или ошибок репликации и отличающие свою ДНК от чужеродных последовательностей.

Системы репарации могут узнавать неверно спаренные или видоизмененные азотистые основания ДНК, а также пропуски нуклеотидов и другие изменения, искажающие структуру двойной спирали. Системы эксцизионной репарации надсекают ДНК вблизи места повреждения, удаляют фрагмент цепи между надрезами и синтезируют новую нуклеотидную последовательность взамен удаленному материалу. Главный путь эксцизионной репарации у *E. coli* обеспечивает система Uvr. Система MutSL участвует в пострепликативной коррекции некомплементарно спаренных оснований в местах встраивания в дочернюю цепь неправильных нуклеотидов. Она функционирует, предпочтительно удаляя основание из той цепи ДНК, которая не метилирована по сайту Dam. Эукариотические гомологи системы MutSL *E. coli* задействованы в репарации неспаренных оснований, порожденных «проскальзыванием» репликации; мутации по этому пути часто сопутствуют определенным типам рака.

Системы рекомбинационной репарации заимствуют информацию из интактного дуплекса и используют ее для починки участка гомологичного дуплекса, ущербного по обеим цепям. Пути RecBC и RecF предваряют действие RecA, чья функция переносчика цепей актуальна при любой бактериальной рекомбинации. Главное предназначение рекомбинационной репарации может заключаться в освобождении застрявших репликативных вилок.

В принципиально другом качестве RecA выступает, когда индуцирует SOS-ответ. Неизвестным способом поврежденная ДНК активирует RecA, и он срабатывает на запуск автокаталитического расщепления белка-репрессора LexA, таким образом снимая блок с экспрессии во многочисленных локусах и индуцируя наработку ферментов-участников путей как эксцизионной, так и рекомбинационной репарации. Гены, контролируемые LexA, содержат операторный SOS-бокс. Некоторые репарационные активности RecA стимулирует напрямую, участием в протеолизе. Расщепление репрессоров лизогенных фагов может побудить фагов к вхождению в литический цикл.

И у про-, и у эукариот системы репарации и транскрипции могут быть взаимосвязаны. Ряд заболеваний человека обусловлен мутациями в генах, кодирующих ферменты репарации, ассоциированные с фактором транскрипции $TF_{II}H$. У них есть гомологи среди генов *RAD* дрожжей, что говорит о филогенетической распространенности этой системы репарации.

Негомологичное соединение концов (NHEJ) представляет собой общий случай починки рваных концов (эукариотической) ДНК. Гетеродимер Ku сводит рваные концы вместе, так что они могут быть склеены (лигированы). Несколько болезней человека обусловлены мутациями компонентов этого пути.

Литература

20.2 Системы репарации корректируют повреждения ДНК

Обзор

Sancar, A., Lindsey-Boltz, L. A., Unsal-Kaçmaz, K., and Linn, S. (2004). Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu. Rev. Biochem.* 73, 39–85.

Wood, R. D., Mitchell, M., Sgouros, J., and Lindahl, T. (2001). Human DNA repair genes. *Science* 291, 1284–1289.

20.4 Способы эксцизионной репарации в клетках млекопитающих

Обзор

Barnes, D. E. and Lindahl, T. (2004). Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu. Rev. Genet.* 38, 445–476.

McCullough, A. K., Dodson, M. L., and Lloyd, R. S. (1999). Initiation of base excision repair: glycosylase mechanisms and structures. *Annu. Rev. Biochem.* 68, 255–285.
Sancar, A., Lindsey-Boltz, L. A., Unsal-Kaçmaz, K., and Linn, S. (2004). Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu. Rev. Biochem.* 73, 39–85.

Исследование

Klungland, A. and Lindahl, T. (1997). Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). *EMBO J.* 16, 3341–3348.
Matsumoto, Y. and Kim, K. (1995). Excision of deoxyribose phosphate residues by DNA polymerase beta during DNA repair. *Science* 269, 699–702.

20.5 Метилазы и гликозилазы вынимают основания из ДНК

Исследование

Aas, P. A., Otterlei, M., Falnes, P. A., Vagbe, C. B., Skorpen, F., Akbari, M., Sundheim, O., Bjoras, M., Slupphaug, G., Seeberg, E., and Krokan, H. E.,(2003). Human and bacterial oxidative demethylases repair alkylation damage in both РНК and DNA. *Nature* 421, 859–863.
Falnes, P. A., Johansen, R. F., and Seeberg, E. (2002). AlkB-mediated oxidative demethylation reverses DNA damage in *E. coli. Nature* 419, 178–182.
Klimasauskas, S., Kumar, S., Roberts, R. J. and Cheng, X. (1994). HhaI methyltransferase flips its target base out of the DNA helix. *Cell* 76, 357–369.
Lau, A. Y., Glassner, B. J., Samson, L. D., and Ellenberger, T. (2000). Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG. *Proc. Natl. Acad. Sci. USA* 97, 13573–13578.
Lau, A. Y., Scherer, O. D., Samson, L., Verdine, G. L., and Ellenberger, T. (1998). Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision. *Cell* 95, 249–258.
Mol, D. D. et al. (1995). Crystal structure and mutational analysis of human uracil-DNA glycosylase: structural basis for specificity and catalysis. *Cell* 80, 869–878.
Park, H. W., Kim, S. T., Sancar, A., and Deisenhofer, J. (1995). Crystal structure of DNA photolyase from *E. coli. Science* 268, 1866–1872.
Savva, R. et al. (1995). The structural basis of specific base-excision repair by uracil-DNA glycosylase. *Nature* 373, 487–493.
Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T., and Sedgwick, B. (2002). Oxidative demethylation by *E. coli* AlkB directly reverts DNA base damage. *Nature* 419, 174–178.
Vassylyev, D. G. et al. (1995). Atomic model of a pyrimidine dimer excision repair enzyme complexed with a DNA substrate: structural basis for damaged DNA recognition. *Cell* 83, 773–782.

20.6 Ошибки репаративного синтеза и фенотипы мутаторов

Исследование

Friedberg, E. C., Feaver, W. J., and Gerlach, V. L. (2000). The many faces of DNA polymerases: strategies for mutagenesis and for mutational avoidance. *Proc. Natl. Acad. Sci. USA* 97, 5681–5683.
Goldsmith, M., Sarov-Blat, L., and Livneh, Z. (2000). Plasmid-encoded MucB protein is a DNA polymerase (pol RI) specialized for lesion bypass in the presence of MucA, RecA, and SSB. *Proc. Natl. Acad. Sci. USA* 97, 11227–11231.
Maor-Shoshani, A., Reuven, N. B., Tomer, G., and Livneh, Z. (2000). Highly mutagenic replication by DNA polymerase V (UmuC) provides a mechanistic basis for SOS untargeted mutagenesis. *Proc. Natl. Acad. Sci. USA* 97, 565–570.
Wagner, J., Gruz, P., Kim, S. R., Yamada, M., Matsui, K., Fuchs, R. P., and Nohmi, T. (1999). The *dinB* gene encodes a novel *E. coli* DNA polymerase, DNA pol IV, involved in mutagenesis. *Mol. Cell* 4, 281–286.

20.7 Контроль направления репарации несовершенных пар оснований

Обзор

Kunkel, T. A., and Erie, D. A. (2005). DNA mismatch repair. *Annu. Rev. Biochem.* 74, 681–710.

Исследование

Strand, M., Prolla, T. A., Liskay, R. M., and Petes, T. D. (1993). Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature* 365, 274–276.

20.8 Системы рекомбинативной репарации *E.coli*

Обзор

West, S. C. (1997). Processing of recombination intermediates by the RuvABC proteins. *Annu. Rev. Genet.* 31, 213–244.

Исследование

Bork, J. M. and Inman, R. B. (2001). The RecOR proteins modulate RecA protein function at 5′ ends of single-stranded DNA. *EMBO J.* 20, 7313–7322.

20.9 Рекомбинация служит важным механизмом исправления ошибок репликации

Обзор

Cox, M. M., Goodman, M. F., Kreuzer, K. N., Sherratt, D. J., Sandler, S. J., and Marians, K. J. (2000). The importance of repairing stalled replication forks. *Nature* 404, 37–41.
McGlynn, P. and Lloyd, R. G. (2002). Recombinational repair and restart of damaged replication forks. *Nat. Rev. Mol. Cell Biol.* 3, 859–870.
Michel, B., Viguera, E., Grompone, G., Seigneur, M., and Bidnenko, V. (2001). Rescue of arrested replication forks by homologous recombination. *Proc. Natl. Acad. Sci. USA* 98, 8181–8188.

Исследование

Courcelle, J. and Hanawalt, P. C. (2003). RecA-dependent recovery of arrested DNA replication forks. *Annu. Rev. Genet.* 37, 611–646.
Kuzminov, A. (2001). Single-strand interruptions in replicating chromosomes cause double-strand breaks. *Proc. Natl. Acad. Sci. USA* 98, 8241–8246.
Rangarajan, S., Woodgate, R., and Goodman, M. F. (1999). A phenotype for enigmatic DNA polymerase II: a pivotal role for pol II in replication restart in UV-irradiated *Escherichia coli. Proc. Natl. Acad. Sci. USA* 96, 9224–9229.

20.10 RecA запускает SOS-систему

Исследование

Tang, M. et al. (1999). UmuD′2C is an error-prone DNA polymerase, *E. coli* pol V. *Proc. Natl. Acad. Sci. USA* 96, 8919–8924.

20.11 Системы репарации эукариотических клеток консервативны

Обзор

Krogh, B. O. and Symington, L. S. (2004). Recombination proteins in yeast. *Annu. Rev. Genet.* 38, 233–271.

Prakash, S. and Prakash, L. (2002). Translesion DNA synthesis in eukaryotes: a one- or two-polymerase affair. *Genes Dev.* 16, 1872–1883.

Sancar, A., Lindsey-Boltz, L. A., Unsal-Kaçmaz, K., and Linn, S. (2004). Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu. Rev. Biochem.* 73, 39–85.

Исследование

Friedberg, E. C., Feaver, W. J., and Gerlach, V. L. (2000). The many faces of DNA polymerases: strategies for mutagenesis and for mutational avoidance. *Proc. Natl. Acad. Sci. USA* 97, 5681–5683.

Johnson, R. E., Prakash, S., and Prakash, L. (1999). Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Pol eta. *Science* 283, 1001–1004.

Rattray, A. J. and Strathern, J. N. (2003). Error-prone DNA polymerases: when making a mistake is the only way to get ahead. *Annu. Rev. Genet.* 37, 31–66.

Reardon, J. T. and Sancar, A. (2003). Recognition and repair of the cyclobutane thymine dimer, a major cause of skin cancers, by the human excision nuclease. *Genes Dev.* 17, 2539–2551.

Wolner, B., van Komen, S., Sung, P., and Peterson, C. L. (2003). Recruitment of the recombinational repair machinery to a DNA double-strand break in yeast. *Mol. Cell* 12, 221–232.

20.12 Неспециализированная система репарации двухцепочечных разрывов

Обзор

D'Amours, D. and Jackson, S. P. (2002). The Mre11 complex: at the crossroads of DNA repair and checkpoint signalling. *Nat. Rev. Mol. Cell Biol.* 3, 317–327.

Исследование

Carney, J. P., Maser, R. S., Olivares, H., Davis, E. M., Le Beau, M., Yates, J. R., Hays, L., Morgan, W. F., and Petrini, J. H. (1998).

The hMre11/hRad50 protein complex and Nijmegen breakage syndrome: linkage of double-strand break repair to the cellular DNA damage response. *Cell* 93, 477–486.

Cary, R. B., Peterson, S. R., Wang, J., Bear, D. G., Bradbury, E. M., and Chen, D. J. (1997). DNA looping by Ku and the DNA-dependent protein kinase. *Proc. Natl. Acad. Sci. USA* 94, 4267–4272.

Ellis, N. A., Groden, J., Ye, T. Z., Straughen, J., Lennon, D. J., Ciocci, S., Proytcheva, M., and German, J. (1995). The Bloom's syndrome gene product is homologous to RecQ helicases. *Cell* 83, 655–666.

Ma, Y., Pannicke, U., Schwarz, K., and Lieber, M. R. (2002). Hairpin opening and overhang processing by an Artemis/DNA-Dependent protein kinase complex in nonhomologous end joining and V(D)J Recombination. *Cell* 108, 781–794.

Ramsden, D. A. and Gellert, M. (1998). Ku protein stimulates DNA end joining by mammalian DNA ligases: a direct role for Ku in repair of DNA double-strand breaks. *EMBO J.* 17, 609–614.

Varon, R. et al. (1998). Nibrin, a novel DNA double-strand break repair protein, is mutated in Nijmegen breakage syndrome. *Cell* 93, 467–476.

Walker, J. R., Corpina, R. A., and Goldberg, J. (2001). Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* 412, 607–614.

Транспозоны

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

21.1 Введение

21.2 Вставочные последовательности — это простейшие модули транспозиции

- Вставочная последовательность представляет собой транспозон с инвертированными концевыми повторами и кодирующий нужный для транспозиции фермент.
- В процессе интеграции транспозона сайт-мишень дуплицируется, образуя два прямых повтора, фланкирующих транспозон.
- Длина прямого повтора составляет от 5 до 9 пн и служит характеристикой разновидности транспозона.

21.3 Композитные транспозоны содержат IS-модули

- Помимо кодирования функций собственного перемещения, транспозоны могут быть носителями других генов.
- У композитных транспозонов есть центральная область, фланкированная IS-элементами.
- Как один, так и оба IS-элемента композитного транспозона в состоянии обеспечить транспозицию.
- Композитный транспозон может переноситься как единое целое, однако любой его активный IS-элемент может переноситься и независимо.

21.4 Транспозиция может идти по репликативному или по нерепликативному механизмам

- Все транспозоны пользуются общей схемой встраивания, согласно которой в ДНК-мишень вносятся смещенные надрезы, к одноцепочечным выступам образовавшихся свободных концов присоединяется транспозон, а бреши заполняются.
- Репликативный либо нерепликативный характер транспозиции обусловлен последовательностью развития событий и конкретным способом связывания транспозона с ДНК-мишенью.

21.5 Транспозоны перестраивают ДНК

- Гомологичная рекомбинация между копиями транспозона вызывает перестройку ДНК хозяина.
- Гомологичная рекомбинация между повторами транспозона ведет к его точному или неточному вырезанию.

21.6 Общие посредники и полупродукты транспозиции

- Транспозиция начинается с образования комплекса переноса цепи, в составе которого транспозон оказывается соединенным с сайтом мишенью одной цепью с каждой стороны.
- Транспозаза Mu формирует комплекс, сближая края последовательности фага Mu; далее проводятся внесение надрезов и реакция переноса цепи.
- Репликативная транспозиция — следствие репликации комплекса, нерепликативная — его репарации.

21.7 Репликативная транспозиция сопровождается образованием коинтеграта

- Репликация в сочетании с переносом цепей порождает коинтеграт — продукт слияния репликонов донора и мишени.
- Коинтеграт содержит две прямых копии транспозона, разделяющих собой последовательности исходных репликонов.
- Рекомбинация между копиями транспозона разделяет коинтеграт на исходные репликоны, причем реципиент получает копию транспозона.
- Реакцию рекомбинации катализирует резольваза, кодируемая транспозоном.

21.8 Нерепликативная транспозиция сопровождается разрывом и воссоединением

- Нерепликативная транспозиция сопровождается разрезанием перекрестной структуры по двум интактным цепям донора и лигированием образовавшихся концов транспозона с оставшимися концами мишени.
- При нерепликативной транспозиции Mu первая пара разрезанных цепей оказывается пришита к мишени до того, как разрезана вторая; у Tn5 и Tn10 все четыре разреза вносятся до пришивания к мишени какого-либо из концов.

21.9 В транспозиции TnA участвуют транспозаза и резольваза

- Репликативная транспозиция TnA требует участия транспозазы (для образования коинтеграта) и резольвазы (для разделения репликонов донора и мишени).
- Резольваза по своему действию напоминает белок Int фага лямбда; по своему типу реакция относится к топоизомеразоподобным сайт-специфичным рекомбинациям с образованием промежуточной ковалентной сшивки между белком и ДНК.

21.10 Контроль транспозиции Tn10 осуществляется различными путями

- Мультикопийное ингибирование снижает вероятность встраивания каждой новой копии транспозона, если другие копии того же транспозона уже встроились в геном.

- Различные молекулярные механизмы влияют на частоту транспозиции.

21.11 Контролирующие элементы вызывают разломы и перестройки в геноме кукурузы

- Явление транспозиции было открыто у кукурузы благодаря разломам хромосом, образующимся при транспозиции мобильных контролирующих элементов.
- Разлом порождает хромосому с центромерой и ломаным концом, а также ацентрический фрагмент (обломок).
- Ацентрический фрагмент теряется во время митоза, это можно заметить по исчезновению доминантных аллелей в гетерозиготе.
- Слияние между ломаными концами двух хромосом порождает дицентрическую хромосому, вовлекаемую в последующие циклы разломов и слияний.
- Цикл разлома-слияния-моста приводит к проявлениям соматической мозаичности.

21.12 Контролирующие элементы образуют семейства транспозонов

- В каждом семействе транспозонов кукурузы есть автономные и неавтономные контролирующие элементы.
- Автономные контролирующие элементы кодируют белки, позволяющие им перемещаться.
- Неавтономные контролирующие элементы содержат мутации, исключающие наличие у них продуктов, способных к катализу транспозиции; все же они могут перемещаться при условии, что автономный элемент из того же семейства обеспечит их необходимыми белками.
- Автономные контролирующие элементы претерпевают смены фаз функциональности, что обусловлено главным образом динамикой метилированности их ДНК.

21.13 Элементы *Spm* влияют на экспрессию генов

- Элемент из семейства *Spm* может влиять на экспрессию генов по месту своего внедрения, если его края связывает белок TnpA.
- Элемент *Spm* инактивируется метилированием.

21.14 Роль мобильных элементов в гибридном дисгенезе

- Особи линии P *Drosophila melanogaster* являются носителями генетических P-элементов, представляющих собой транспозоны.
- В результате скрещивания самца P с самкой M транспозиция P-элементов активируется.
- Внедрение P-элементов в новые сайты инактивирует множество генов в геноме потомства от скрещивания самца P с самкой M, приводя к бесплодию.

21.15 P-элементы могут быть активированы в клетках зародышевого пути

- При скрещивании самца P и самки M P-элементы в клетках зародышевого пути их потомства активно перемещаются по геному, поскольку при удалении одного интрона в ходе сплайсинга, специфичного для клеток зародышевого пути, образуется последовательность, кодирующая транспозазу.
- Другой белковый продукт P-элемента — репрессор транспозиции, синтезируется во всех клетках, и его очень много в цитоплазме яйцеклеток самки P.
- Отсутствие признаков дисгенеза у гибрида между самцом M и самкой P объясняется присутствием репрессора в цитоплазме оплодотворенной яйцеклетки.

21.16 Резюме

21.1 Введение

Геномы эволюционируют, приобретая новые последовательности и перестраивая имеющиеся.

Внезапное внесение новых последовательностей обусловлено способностью векторов к переносу информации между геномами. Внехромосомные элементы опосредуют горизонтальный поток информации, перенося фрагменты генетического материала (обычно довольно короткие). У бактерий плазмиды переносятся при конъюгации, в то время как фаги распространяются путем инфекции (см. гл. 14 «Стратегии бактериофагов»). Плазмида или фаг при переносе собственного репликона могут передавать и гены хозяина. Прямой перенос ДНК свойственен некоторым бактериям: они обмениваются материалом друг с другом путем трансформации (см. разд. 1.2 «ДНК — это генетический материал бактерий»). Некоторые вирусы эукариот (в особенности ретровирусы) могут переносить генетическую информацию в ходе инфекционного цикла (см. разд. 22.6 «Ретровирусы могут переносить клеточные последовательности»).

Перестройки обеспечиваются процессами, проходящими внутри генома. Две главных причины геномных перестроек резюмирует РИС. 21.1.



РИС. 21.1 Изменения нуклеотидной последовательности генома чаще всего связаны с перемещением транспозона на новое место. События такого рода могут сильно влиять на экспрессию генов. Также геномные перестройки могут являться результатом неравного кроссинговера между похожими последовательностями, которыми также могут оказаться копии транспозона

Неравная рекомбинация — результат неверной состыковки материала клеточными системами гомологичной рекомбинации. Нереципрокная рекомбинация приводит к дупликации или перестройке локусов (см. разд. 6.7 «Неравный кроссинговер приводит к реорганизации целых генетических кластеров»). Дупликация внутри генома служит главным источником большинства новых последовательностей, в том смысле что одна копия последовательности будет обеспечивать исходную функцию, а вторая сможет эволюционировать в новую функцию. Кроме того, даже между близкородственными геномами существуют различия, заключающиеся в полиморфных вариациях — эффектах событий рекомбинации. В разд. 6.14 «Мини-сателлиты можно использовать при генетическом картировании» мы видели, что рекомбинация между мини-сателлитами переиначивает их размеры, так что каждый индивидуальный геном неповторимо отличается от других.

Еще одна причина геномной изменчивости — существование **мобильных элементов**, или **транспозонов**. Они представляют собой дискретные (обособленные) подвижные последовательности, способные перемещаться по разным местоположениям в пределах генома. Транспозон (в этом его отличительная особенность) не принимает, подобно фагу или плазмиде, форму независимого элемента, а выполняет переход непосредственно с одного геномного сайта на другой. В отличие от большинства других процессов, лежащих в основе реструктуризации генома, транспозиция не опирается на какое-либо сходство последовательностей сайтов донора и реципиента. В перемещениях своего, а иногда и дополнительного генетического материала, транспозоны ограничены пространством одного единственного генома. Они — внутригеномные аналоги векторов, способных транспортировать материал из одного генома в другой.

Транспозоны делят на два класса. В этой главе рассмотрены транспозоны, существующие как кодирующие белки последовательности, способные напрямую манипулировать ДНК, чтобы распространять себя внутри генома. Гл. 22 «Ретровирусы и ретропозоны» посвящена транспозонам, родственным ретровирусам: их подвижность проистекает от способности делать ДНК-копии собственных РНК-транскриптов; эти копии интегрируют в новые сайты генома.

Транспозоны, мобилизуемые через ДНК, обнаружены и у про-, и у эукариот. Каждый бактериальный транспозон несет в себе ген(ы), кодирующие энзиматические активности, необходимые ему для транспозиции, хотя для перемещения могут потребоваться вспомогательные функции, закодированные в геноме, в который они встроены (такие как ДНК-полимераза или ДНК-гираза). Подобные системы есть и у эукариот, правда, их энзиматические функции изучены не так хорошо. В геноме могут одновременно присутствовать и функциональные, и нефункциональные (дефектные) элементы. Нередко бывает, что дефектных мобильных элементов в эукариотическом геноме — большинство: они потеряли способность перемещаться самостоятельно, но ферменты, продуцируемые функциональными транспозонами, все еще могут распознавать их как субстрат для перемещения. Эукариотический геном содержит огромное число и разнообразие транспозонов. В геноме мухи их более 50 разновидностей, общим числом в несколько сотен индивидуальных элементов.

Мобильные элементы могут прямо или косвенно способствовать перестройкам генома:

- Само событие транспозиции может вызвать делецию или инверсию фрагмента ДНК или повлечь за собой перемещение последовательности фрагмента генома хозяина в новое положение.
- Будучи «портативными участками гомологии», транспозоны служат субстратом для клеточной системы рекомбинации; две копии транспозона в разных положениях (даже на разных хромосомах) могут сработать в качестве сайтов реципрокной рекомбинации. Это чревато делециями, вставками, инверсиями и транслокациями.

С точки зрения естественного отбора, транспозон с его скачкообразной активностью — мишень довольно расплывчатая. Логично полагать мобильные элементы (во всяком случае некоторые из них) ни выигрышными, ни проигрышными для фенотипа, но способными организоваться в эгоистичную ДНК — ДНК, занимающуюся только собственным распространением В самом деле, рассматривая транспозицию как явление, стоящее особняком от других клеточных систем рекомбинации, мы тем самым принимаем точку зрения на транспозон как независимую составляющую генома.

Взаимоотношения между транспозоном и геномом напоминают взаимоотношения между паразитом и его хозяином. По-видимому, распространение элемента посредством транспозиции уравновешено вредом, причиняемым клетке, когда транспозиция инактивирует важный ген, или когда обилие транспозонов становится обременительным для клеточных систем. Все же нельзя забывать и о том, что событие транспозиции, повлекшее селективное преимущество (например, вследствие геномной перестройки), оборачивается приоритетным выживанием генома, несущего активный транспозон.

21.2 Вставочные последовательности — это простейшие модули транспозиции

Ключевые понятия:

- Вставочная последовательность представляет собой транспозон с инвертированными концевыми повторами и кодирующий нужный для транспозиции фермент.
- В процессе интеграции транспозона сайт-мишень дуплицируется, образуя два прямых повтора, фланкирующих транспозон.
- Длина прямого повтора составляет от 5 до 9 пн и служит характеристикой разновидности транспозона.

Мобильные элементы на молекулярном уровне были впервые идентифицированы в форме вставок, спонтанно

возникающих в бактериальных оперонах. Такая вставка препятствует транскрипции и/или трансляции оккупируемого ею гена. К настоящему моменту описано множество разновидностей мобильных элементов подобного рода.

Простейшие транспозоны называют **вставочными последовательностями**; такое название отражает историю их обнаружения. Каждый тип вставочной последовательности обозначают номером с префиксом **IS**. Те, что были обнаружены в самом начале, получили номера от IS1 до IS4; более поздние разновидности получали номера, отражающие обстоятельства эксперимента, приведшего к их идентификации (номера эти не связаны с общим числом разновидностей, которые удалось изолировать!).

IS-элементы — естественные составляющие бактериальных хромосом и плазмид. Стандартный штамм *E. coli* содержит на клетку по нескольку (<10) копий самых распространенных разновидностей IS-элементов. Для обозначения местонахождения вставки в определенном сайте используют двойное двоеточие; так, выражение λ::IS1 обозначает IS1-элемент, вставленный в фаг λ.

IS-элементы — автономные единицы, кодирующие только белки, необходимые для поддержки собственной транспозиции. Каждая разновидность уникальна по последовательности, но в организации всех IS-элементов есть ряд характерных общих черт. Структура типичного транспозона до и после интеграции (внедрения) в сайт-мишень показана на РИС. 21.2, который, помимо этого, суммирует детали строения наиболее часто встречающихся IS-элементов.

IS-элемент ограничивают короткие **инвертированные концевые повторы**; обычно левая копия повтора не совсем идентична правой. Как показано на рисунке, и в буквальном понимании, инвертированный концевой повтор представляет собой элемент, встречающийся с любого фланга ДНК и читаемый в одном и том же порядке (правда, на *разных* цепях ДНК). При транспозиции IS-элемента последовательность ДНК хозяина в сайте внедрения удваивается. Устройство этой дупликации выявляется сравнением последовательностей сайта-мишени до и после внедрения. На рис. 21.2 показано, что ДНК IS-вставки всегда фланкирована очень короткими **прямыми повторами**, их одинаковые последовательности находятся в одной и той же ориентации. До внедрения последовательность сайта-мишени соответствует одной копии повтора. В примере, изображенном на рисунке, сайт-мишень представлен пентануклеотидом $\begin{matrix}\mathrm{ATGCA}\\ \mathrm{TACGT}\end{matrix}$ и комплементарной последовательностью второй цепи ДНК. После транспозиции с каждой стороны от вставки возникает по одной копии этого сайта.

Для отдельно взятого IS-элемента нуклеотидная последовательность прямого повтора может изменяться с каждым актом транспозиции, но длина его для каждой разновидности IS-элемента постоянна. Это отражает механизм переноса; самая частая длина прямого повтора равна 9 пн.

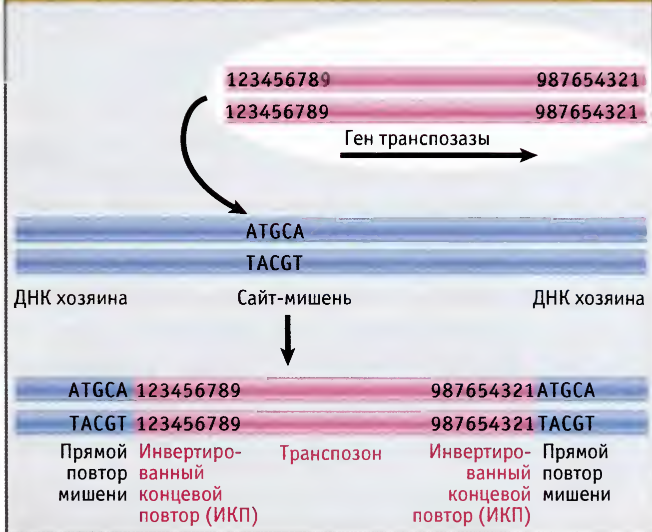
Таким образом, края IS-элемента представляют собой инвертированные концевые повторы, а примыкающие к ним края ДНК хозяина представляют собой короткие прямые повторы. Это — главный диагностический признак, по которому принято идентифицировать транспозоны в составе последовательностей ДНК. То есть вычленение такой структуры у какой-либо последовательности ДНК указывает на происхождение данной последовательности от события транспозиции.

Большинство IS-элементов вставляется по разнообразным сайтам ДНК хозяина. Однако некоторые разновидности отдают предпочтение (различной степени) конкретным горячим точкам.

Итак, края транспозона обозначены инвертированными повторами. Опознание краев — общий момент при переносе любого мобильного элемента. *Цис*-действующие мутации, препятствующие транспозиции, сосредоточены как раз в краях элемента. Эти края опознает специальный белок — **транспозаза**.

Все IS-элементы, за исключением IS1, содержат единственную кодирующую область, начинающуюся сразу вглубь от внутреннего конца одного инвертиро-

Транспозон содержит инвертированные повторы и находится в окружении прямых повторов



Транспозон	Длина повторов мишени, пн	Длина ИКП, пн	Общая длина, пн	Выбор мишени
IS1	9	23	768	случайный
IS2	5	41	1327	горячие точки
IS4	11–13	18	1428	$AAAN_{20}TTT$
IS5	4	16	1195	горячие точки
IS10R	9	22	1329	NGCTNAGCN
IS50R	9	9	1531	горячие точки
IS903	9	18	1057	случайный

РИС. 21.2 Транспозоны содержат инвертированные концевые повторы и фланкированы прямыми повторами сайта-мишени. В примере на рисунке сайт-мишень имеет длину 5 пн, в то время как края транспозона представляют собой инвертированные повторы длиной 9 пн. На рисунке пары оснований в инвертированных повторах пронумерованы от 1 до 9

ванного повтора и заканчивается вплотную к инвертированному повтору на другом краю (или даже внутри него). Эта область кодирует транспозазу. У IS1-элемента — более сложная организация, с двумя отдельными рамками считывания; использование при продукции транспозазы их обеих достигается благодаря сдвигу рамки в процессе трансляции.

Частота транспозиций варьирует, в среднем составляя от 10^{-4} до 10^{-3} на элемент за поколение. Внедрение в отдельно взятую мишень по частоте сопоставимо со спонтанным мутагенезом (от 10^{-7} до 10^{-5} за поколение). Реверсия IS-элементов (точное вырезание и переворот на месте), как правило, редка и может составить от 10^{-10} до 10^{-6} события на элемент за поколение, то есть не более 10^{-3} части всех транспозиций.

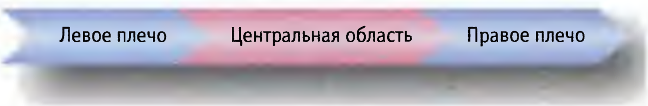
21.3 Композитные транспозоны содержат IS-модули

Ключевые понятия:

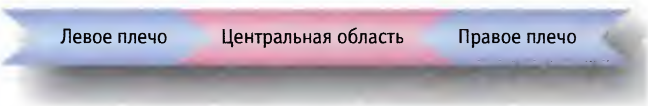
- Помимо кодирования функций собственного перемещения, транспозоны могут быть носителями других генов.
- У композитных транспозонов есть центральная область, фланкированная IS-элементами.
- Как один, так и оба IS-элемента композитного транспозона в состоянии обеспечить транспозицию.
- Композитный транспозон может переноситься как единое целое, однако любой его активный IS-элемент может переноситься и независимо.

Некоторые транспозоны, помимо кодирования энзиматической базы транспозиции, несут какие-либо другие генетические маркеры — например, обусловливающие устойчивость к антибиотикам. Таким транспозонам присваивают обозначение Tn и номер. Один из классов крупных транспозонов называют **композитными элементами** за то, что их центральная область, несущая маркеры устойчивости к антибиотикам, окаймлена плечами из IS-элементов.

Плечи (одинаковые или почти одинаковые) могут быть в одной ориентации, но чаще инвертированы. Композитный транспозон с плечами в виде прямых повторов имеет структуру



Если плечи представляют собой инвертированные повторы, структура будет такой:

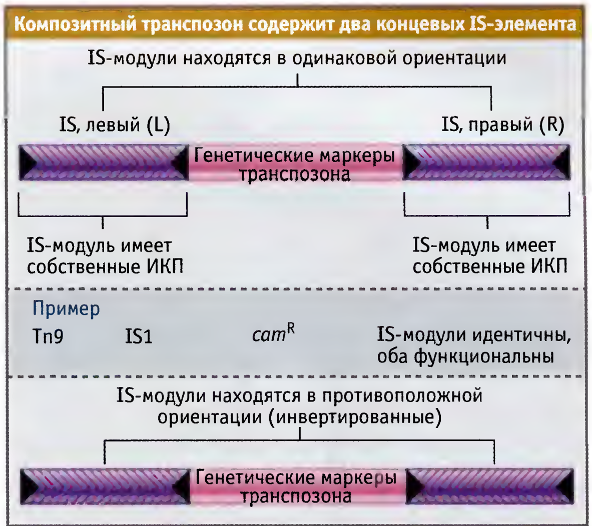


Стрелками обозначена ориентация плеч, которые считаются левым и правым в соответствии с (произвольной) ориентацией карты транспозона слева направо. Более подробно структура композитного транспозона проиллюстрирована на рис. 21.3, который, кроме того, обобщает свойства наиболее распространенных композитных транспозонов.

Обычно каждое IS-плечо несет по краям инвертированные концевые повторы; таким образом, композитный транспозон как целое окаймляют те же короткие инвертированные концевые повторы.

В некоторых случаях IS-модули композитного транспозона идентичны, например плечи Tn9 представляют собой прямо повторенный IS1, а плечи Tn903 — инвертированные копии IS903. В других случаях, как у Tn10 или Tn5, модули очень похожи, но не идентичны, что позволяет различать левые и правые модули у Tn10 или Tn5.

Функциональный IS-модуль может обеспечить транспозицию как одного себя, так и композитного транспозона целиком. Когда модули идентичны (как у Tn9 или Tn903), поддержку перемещению транспозона, по всей видимости, способен оказать каждый из них. Если модули отличаются, могут быть разными и их функциональные возможности, так что транспозиция может (всецело или главным образом) зависеть от одного из модулей (как у Tn10 или Tn5).



Пример	Левый край	Маркеры	Правый край
Tn903	IS903	kan^R	Оба IS-модуля функциональны
Tn10	IS10L не функционален	tet^R	IS10R функционален
Tn5	IS50L не функционален	kan^R	IS50R функционален

21.3 У композитного транспозона есть центральная область, содержащая генетические маркеры (например, гены устойчивости к антибиотикам). Центральная область фланкирована IS-модулями (они также являются частью транспозона). Каждый модуль содержит короткие ИКП. Если правый и левый модули находятся в противоположной ориентации (то есть являются инвертированными), их дистальные ИКП находятся в одинаковой ориентации: получается, что композитный транспозон содержит короткие прямые концевые повторы

Предполагают, что эволюционно композитные транспозоны возникли в результате ассоциации двух исходно независимых IS-модулей с центральной областью. Такая ситуация могла сложиться, когда при транспозиции IS-элемента сайт-реципиент (мишень) был расположен близко к сайту-донору. Оказавшись в составе композитного элемента, два идентичных IS-модуля могли таковыми и остаться или же дивергировать. Способность IS-модуля в одиночку обеспечить перемещение всего композитного элемента объясняет отсутствие давления естественного отбора, которое заставляло бы оба модуля оставаться транспозиционно-активными.

Какими же обстоятельствами навязан перенос всего композитного элемента, а не отдельного IS-модуля? Этот вопрос особенно актуален, если оба модуля функциональны. Например, у Tn9 каждый IS1-модуль активен, по всей видимости, и сам по себе, и в составе транспозона. Почему композитный элемент сохраняет целостность и вставочные последовательности плеч не перемещаются индивидуально?

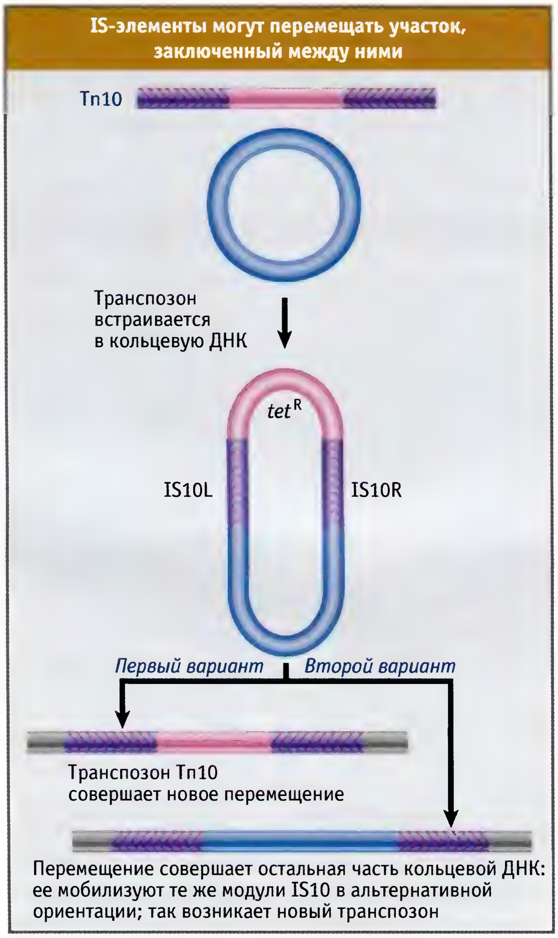


РИС. 21.4 Два модуля IS10 работают как организаторы композитного транспозона и способны перенести в новое место последовательность ДНК, которая их разделяет. Если, например, Tn10 является частью небольшой кольцевой молекулы, модули IS10 способны привести в движение как сам Tn10, так и противолежащую часть кольца

На самом деле два IS-элемента могут переместить любую заключенную между ними последовательность, равно как и свою. На РИС. 21.4 показано, что, если Tn10 пребывает в составе кольцевого репликона, то два его IS10-модуля можно рассматривать не только как плечи центральной области Tn10 с ее геном *tet*^{R} (устойчивости к тетрациклину), но и как фланги противоположного сегмента кольца. Значит, окончательным объектом транспозиции может оказаться как Tn10 (тогда событие будет маркировано движением *tet*^{R}), так и новый «вывернутый» транспозон с альтернативной центральной областью.

Заметим, что у этих транспозонов, исходного и инвертированного, IS10-модули имеют противоположные ориентации относительно центральной области, и это, очевидно, не влияет на их функциональность. Частота событий транспозиции для композитного элемента находится в обратной зависимости от расстояния между модулями (чем дальше, тем реже). Такая зависимость — фактор ограничения размера обыкновенных композитных транспозонов.

Главной силой, подкрепляющей перенос композитных транспозонов, служит отбор по маркерам центральной области. Модуль IS10 свободен перемещаться по собственному усмотрению и в одиночку приходит в движение на порядок чаще, чем Tn10 целиком. Tn10 связан отбором с *tet*^{R}, поэтому в селективных условиях относительная частота транспозиции Tn10 многократно увеличивается.

IS-элементы кодируют транспозазы, ответственные и за подготовку сайтов-мишеней, и за опознание краев транспозона. Чтобы послужить субстратом транспозиции, транспозону не нужно ничего, кроме собственных краев.

21.4 Транспозиция может идти по репликативному или по нерепликативному механизмам

Ключевые понятия:

- Все транспозоны пользуются общей схемой встраивания, согласно которой в ДНК-мишень вносятся смещенные надрезы, к одноцепочечным выступам образовавшихся свободных концов присоединяется транспозон, а бреши заполняются.
- Репликативный либо нерепликативный характер транспозиции обусловлен последовательностью развития событий и конкретным способом связывания транспозона с ДНК-мишенью.

Внедрение (интеграцию) транспозона в новый сайт иллюстрирует РИС. 21.5. Процесс состоит из этапов внесения смещенных надрезов в ДНК-мишень, пришивания транспозона к одноцепочечным выступам краев разреза, заполнения брешей. Порождение и последующее заполнение одноцепочечных концов объясняет наличие прямых повторов ДНК-мишени по обе стороны от встроившегося элемента.

Длина прямых повторов диктуется величиной смещения между надрезами в двух цепях мишени на начальном этапе интеграции. Эта величина отображает геометрию фермента (транспозазы) участвующего в разрезании.

Использование одноцепочечных концевых выступов — общий момент для любого из способов транспозиции; всего же, по механизму перемещения элемента, различают три типа транспозиции:

- При **репликативной траспозиции** элемент дуплицируется, и объект транспортировки представляет собой копию исходного элемента. РИС. 21.6 демонстрирует результат такой транспозиции. Копирование транспозона — это неотъемлемая часть его перемещения. Одна копия остается в исходном сайте (доноре), а другая копия внедряется в новый. Такая транспозиция сопровождается увеличением числа копий элемента. В репликативной транспозиции участвует два типа энзиматической активности: транспозаза действует на границы исходного транспозона, а **резольваза** действует на дуплицированный материал. Транспозоны семейства TnA перемещаются по геному исключительно репликативным способом (см. разд. 21.7 «Репликативная транспозиция сопровождается образованием коинтеграта»).
- При **нерепликативной транспозиции** элемент физически перемещается непосредственно с одного сайта на другой. При этом материал ДНК в составе транспозона не претерпевает изменений, т. е. процесс консервативен. Вставочные последовательности и некоторые композитные элементы (например, Tn10 и Tn5) делают это через полное высвобождение из фланкирующей ДНК сайта-донора. Такого рода механизм, показанный на РИС. 21.7, требует участия одной лишь транспозазы. Другой механизм употребляет связь между ДНК донора и мишени, в нем есть этапы, общие с процессом репликативной транспозиции (см. разд. 21.6 «Общие посредники и полупродукты транспозиции»). Оба механизма нерепликативной транспозиции приводят к исчезновению элемента из сайта-донора с приходом в сайт-мишень. Какова дальнейшая судьба молекулы-донора? Для его выживания необходимо распознавание двухцепочечного разрыва системами репарации хозяина.
- **Консервативная транспозиция** описывает другое нерепликативное событие: вырезание элемента из сайта-донора и внедрение в мишень с сохранением *всех* межнуклеотидных связей (см. РИС. 21.8). Это в точности напоминает механизм интеграции лямбда, который обсуждался в разд. 19.16 «Специализированная рекомбинация задействует специфичные сайты», а транспозазы таких элементов родственны интегразе λ и членам ее семейства. Мобильные элементы, использующие такой механизм, велики и могут опосредовать перенос не только самих себя, но и ДНК донора из одной бактерии в другую. Сначала такие элементы классифицировали как транспозоны, но более правильно считать их эписомами.

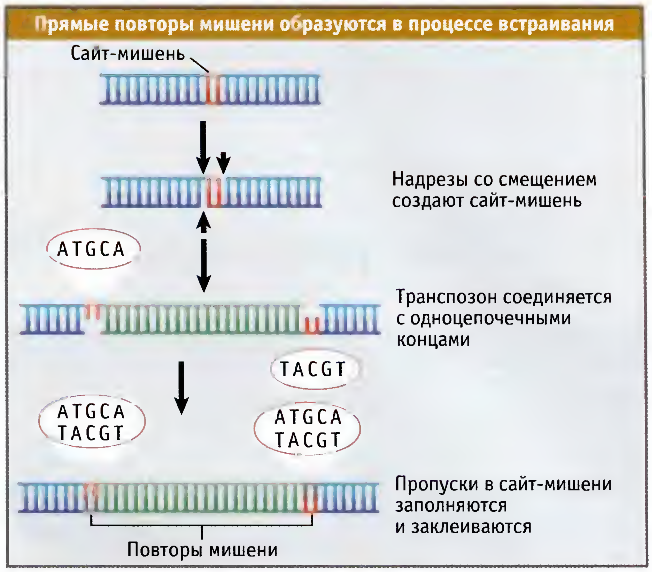


РИС. 21.5 Прямые повторы в ДНК, фланкирующей транспозон, возникают благодаря внесению смещенных надрезов с образованием идентичных концевых выступов, к которым присоединяются концы транспозона при его встраивании

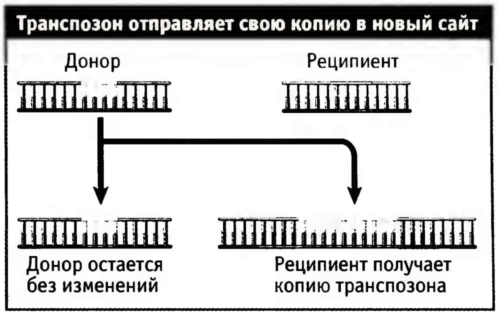


РИС. 21.6 При репликативной транспозиции образуется копия транспозона, которая встраивается в сайт-реципиент. Сайт-донор остается без изменений, так что в результате оба сайта содержат по копии транспозона

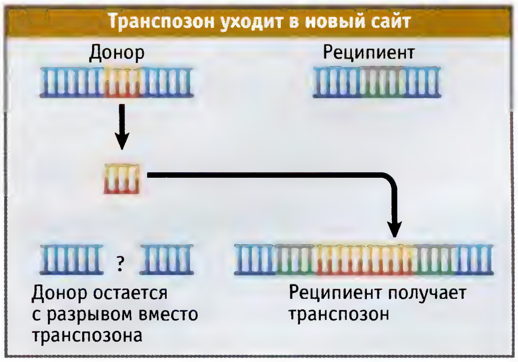


РИС. 21.7 При нерепликативной транспозиции транспозон совершает физическое перемещение из сайта-донора в сайт-реципиент. Двухцепочечный разрыв, оставленный в сайте-доноре, нуждается в репарации

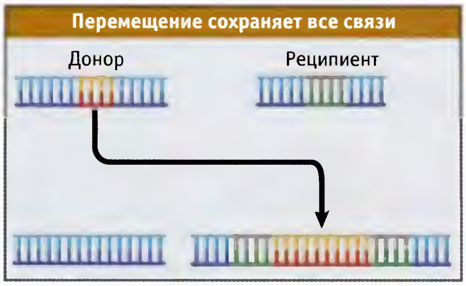


РИС. 21.8 При консервативной транспозиции (например, интеграции и эксцизии лямбда) транспозон совершает физическое перемещение без потери межнуклеотидных связей

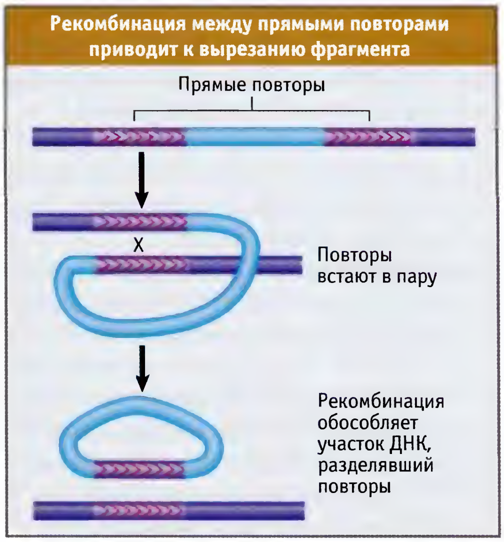


РИС. 21.9 Реципрокная рекомбинация между прямыми повторами в ДНК эквивалентна вырезанию материала между этими повторами. Каждый продукт рекомбинации имеет одну копию повтора

Некоторые транспозоны пользуются только одним механизмом транспозиции, другие способны использовать разные пути. Элементы IS1 и IS903 пользуются и репликативным, и нерепликативным путями; подробно описана способность фага Mu обращаться к тому и другому механизму, исходя из общего для них полупродукта (см. разд. 21.6 «Общие посредники и полупродукты транспозиции»). Одни и те же базовые типы реакций обусловливают любой акт транспозиции. Транспозон и ДНК-донор разъединяются реакциями разрезания, после чего открытые концы транспозонов соединяются с ДНК-мишенью реакциями переноса по методу трансэтерификации: 3'-OH конец цепи транспозона напрямую атакует ДНК-мишень. Эти реакции имеют место внутри нуклеопротеинового комплекса, содержащего необходимые ферменты и оба конца транспозона. Транспозоны также классифицируются по тому, опознает система сайт-мишень до разъединения мобильного элемента с донором или же после.

Выбор сайта-мишени осуществляется транспозазой. В некоторых случаях мишень выбирается практически случайным образом. В других случаях выбор определяется наличием консенсуса или другим свойством ДНК. Этим свойством может быть вторичная структура ДНК (например, изгиб) или ДНК-белковый комплекс. Последний вариант может приводить к внедрению транспозона в специфический промотор (так Ty1 или Ty3 у дрожжей «выбирают» промоторы polIII), в неактивные участки хромосом или в реплицирующуюся ДНК.

21.5 Транспозоны перестраивают ДНК

Ключевые понятия:

- Гомологичная рекомбинация между копиями транспозона вызывает перестройку ДНК хозяина.
- Гомологичная рекомбинация между повторами транспозона ведет к его точному или неточному вырезанию.

Помимо «простой» молекулярной транспозиции — смены местонахождения вставки, — транспозоны могут соучаствовать в других типах перестроек ДНК. Иногда ДНК перестраивается из-за взаимодействия между элементами множества идентичных копий транспозона. Другие перестройки возникают как результат альтернативных исходов транспозиции, служа ключом к разгадке ее механизма.

Встраивание второй копии транспозона поблизости от первой чревато перестройкой ДНК хозяина. Клеточные системы могут подвергнуть пару копий транспозона реципрокной рекомбинации, последствия которой зависят от взаимной ориентации копий.

РИС. 21.9 иллюстрирует общее топологическое правило, согласно которому рекомбинация между парой прямых повторов вырезает материал, распознающий их. Промежуточная область оказывается исключенной из хромосомы вместе с одной (рекомбинантной) копией транспозона в составе кольцевой ДНК (которую клетка теряет). В хромосоме остается единственная копия прямо повторявшегося транспозона. Так, рекомбинация между прямо повторяющимися IS1-модулями транспозона Tn9 приведет к замене всего композитного транспозона одиночным IS1-модулем.

Таким образом, **делеция** смежной с транспозоном последовательности ДНК может быть результатом двухступенчатого процесса. Сначала транспозиция порождает прямую копию элемента, затем между копиями происходит рекомбинация. Однако большинство делеций материала ДНК, происходящих в непосредственной близости от транспозонов, своим возникновением обязаны вариабельности механизма самого события транспозиции.

РИС. 21.10 изображает топологические последствия реципрокной рекомбинации между парой инвертированных (развернутых на 180°) копий транспозона. Область между копиями претерпевает инверсию, а сами они сохраняют свое место в ДНК хозяина — вместе с возможностью и дальше способствовать инверсиям

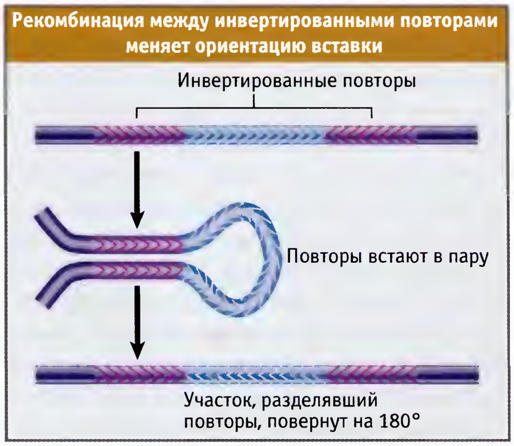


РИС. 21.10 Реципрокная рекомбинация между инвертированными повторами переворачивает материал, заключенный между повторами, относительно геномного окружения

прмежуточной области. Композитный транспозон, чьи краевые модули взаимно инвертированы, представляет собой стабильный компонент генома, хотя из-за рекомбинации между модулями ориентация центральной области относительно хромосомного окружения может (время от времени) меняться.

Эксцизия не поддерживается изолированными транспозонами, но может быть вызвана бактериальными ферментами, распознающими гомологичные участки в составе транспозона. Это важно для клетки, поскольку устрата транспозона иногда помогает восстановить клеточную функцию, нарушенную его внедрением. При **точной эксцизии** удаляется транспозон плюс одна копия последовательности, дуплицировавшейся при его внедрении (то есть один из коротких прямых повторов, фланкирующих транспозон). Это бывает редко; для Tn5 и Tn10 частоты составляют ~10^{-6} и 10^{-9} соответственно. Вероятно, процесс задействует рекомбинацию между короткими прямыми повторами по 9 пн.

Неточная эксцизия оставляет после себя часть транспозона. После такой эксцизии остаточный фрагмент транспозона может препятствовать реактивации гена-мишени, однако если транспозон имел полярный эффект в отношении смежных генов, то эффект этот при неточной эксцизии, как правило, снимается, что влечет изменение фенотипа. Неточная эксцизия Tn10 в обычных условиях происходит с частотой ~10^{-6}. Процесс задействует рекомбинацию между 24-пн фрагментами IS10-модулей. Это последовательности инвертированных повторов по краям IS10-модулей, которые образуют пары прямых повторов с такими же последовательностями из другого IS10-модуля композитного транспозона.

Возможно, неточная эксцизия случается чаще точной из-за того, что опирается на использование более протяженных прямых повторов (24 пн против 9 пн, соответственно, для Tn10). Оба процесса не полагаются на кодируемые транспозоном энзиматические функции, хотя доподлинно их механизмы *не известны*. Во всяком случае, эксцизия транспозонов RecA-независима; возможно, она обусловлена каким-то клеточным механизмом порождения спонтанных делеций между близко расположенными повторяющимися последовательностями.

21.6 Общие посредники и полупродукты транспозиции

Ключевые понятия:

- Транспозиция начинается с образования комплекса переноса цепи, в составе которого транспозон оказывается соединенным с сайтом мишенью одной цепью с каждой стороны.
- Транспозаза Mu формирует комплекс, сближая края последовательности фага Mu; далее проводятся внесение надрезов и реакция переноса цепи.
- Репликативная транспозиция — следствие репликации комплекса, нерепликативная — его репарации.

Механизмы переноса многих мобильных ДНК-элементов — IS-элементов, про- и эукариотических транспозонов, бактериофага Mu — фундаментально похожи. Тем же принципам подчинены и механизмы внедрения ДНК-копии ретровирусной РНК (см. разд. 22.2 «Жизненный цикл ретровируса включает в себя стадии, подобные транспозиции»), и первые этапы рекомбинации иммуноглобулиновых генов (см. разд. 23.10 «Белки RAG катализируют разрыв и воссоединение ДНК»).

Транспозиция начинается соединением транспозона с мишенью. РИС. 21.11 показывает, как транспозон оказывается надрезанным с обоих концов, а сайт-мишень — по обеим цепям. Свободные концы надрезов соединяются крест-накрест — транспозон с мишенью оказываются ковалентно сшитыми. Края транспозона при этом сближены; чтобы было проще следить за внесением надрезов в ДНК, этап сближения (синапсиса) краев транспозона изображен, как более поздний, но в действительности он предшествует внесению надрезов.

Впервые процесс был охарактеризован для фага Mu, который использует две разновидности транспозиции. Инфицируя клетку-хозяина, Mu интегрирует в геном посредством нерепликативной транспозиции; перед выходом в литический цикл число копий преумножается репликативной транспозицией. Оба варианта стартуют одной и той же реакцией между транспозоном и мишенью, но последующие реакции у них разные.

Первоначальные манипуляции с фаговой ДНК проделывает транспозаза MuA. По три сайта связывания MuA, с консенсусом в 22 пн, расположены по краям ДНК Mu: L1, L2, L3 с левого края; R1, R2, R3 с правого края. С каждым из них может связаться мономер MuA. Также MuA может связаться с внутренним сайтом в геноме фага. Связывание MuA по обоим краям и во внутреннем сайте рождает комплекс. Роль внутреннего сайта не совсем ясна; похоже, что он необходим для образования комплекса, но не для внесения надрезов и последующих этапов.

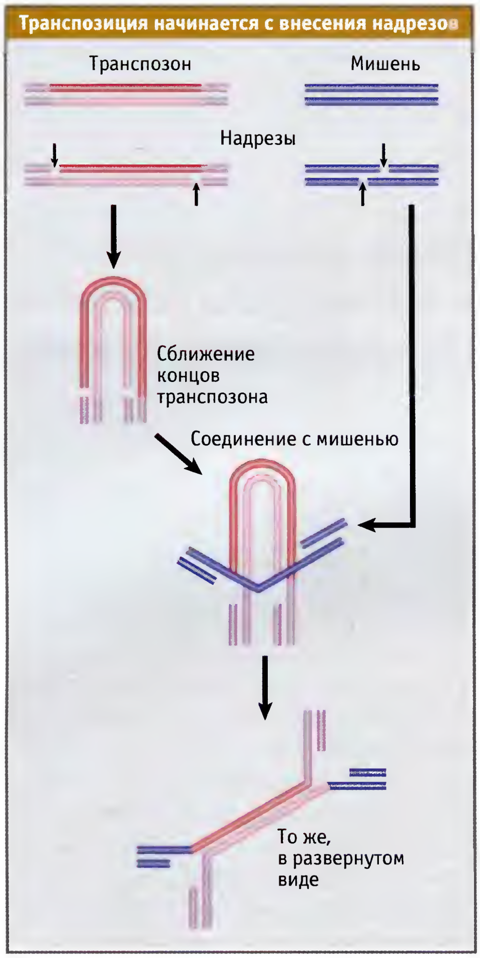


РИС. 21.11 Транспозиция начинается с внесения пары надрезов по концам транспозона и пары надрезов в сайт-мишень. Полученные свободные концы ДНК объединяются в комплекс переноса цепей

Соединение ДНК фага-транспозона Mu с сайтом мишенью проходит в три стадии — они изображены на РИС. 21.12. В соединении участвуют четыре самых крайних сайта связывания MuA: L1, L2, R2 и R1. Субъединицы MuA, которые с ними связываются, образуют тетрамер, благодаря этому края транспозона оказываются сближены. Тетрамер координирует участие концов Mu-ДНК в реакции. Для манипуляций с ДНК тетрамер MuA располагает двумя энзиматически-активными центрами, устройство которых вынуждает субъединицы этой транспозазы действовать в *транс*-манере. Как мы помним, у каждой субъединицы белковый сайт связывания консенсуса ассоциирован с ДНК-сайтом L1, L2, R1 или R2. Энзиматически-активный центр MuA надрезает цепи ДНК Mu по внешним границам сайтов L1 и R1. В силу стереохимических ограничений, он не в состоянии надрезать участок ДНК, смежный с тем, что связан с его же субъединицей. Поэтому мономеры MuA, ассоциированные с консенсусом, режут не ту ДНК, на которой сидят. Один из связанных с левым краем мономеров надрезает правый край, и наоборот. Какие именно два мономера — L1 или L2, R1 или R2 — энзиматически-активны на этом этапе реакции, неизвестно. Перенос цепей тоже осуществляется в *транс*-манере: мономер MuA, связанный с ДНК в L1, работает переносчиком цепи ДНК сайта R1, и наоборот. Может статься, что последовательные этапы надрезания и переноса цепи ДНК по одному краю фага-транспозона катализируются двумя разными мономерами MuA, сидящими на другом краю.

Реакции способствует еще один белок — MuB, влияющий на выбор сайта-мишени. Предпочтение оказывается мишеням, удаленным от сайта-донора на 10–15 тпн. Причина в том, что когда MuB связывается с комплексом MuA-ДНК, то под действием MuA MuB гидролизует АТФ, после чего MuB уходит. А между тем MuB служит очень важным звеном: связавшись с ДНК-мишенью, он стимулирует рекомбинативную активность MuA. Но доминирующее присутствие MuA в окрестности донора блокирует доступ MuB к ДНК, создавая эффект «иммунности мишени». В чистом виде этот эффект демонстрируется экспериментом в системе *in vitro*, состоящей из двух плазмид (Mu-содержащего донора и Mu-пустой мишени), белков MuA и MuB, белка HU *E. coli*, ионов Mg^{2+} и

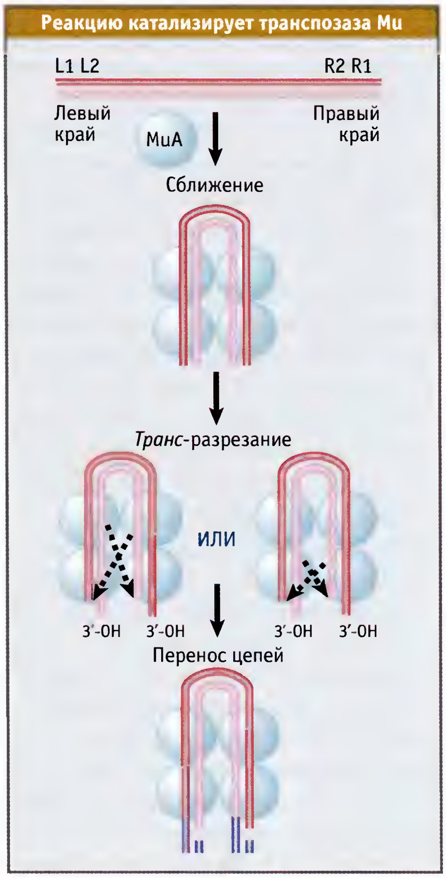


РИС. 21.12 Транспозиция ДНК фага Mu проходит в три равновесных этапа. Субъединицы транспозазы MuA образуют тетрамер, в котором сближаются концы ДНК Mu. Затем транспозаза делает надрезы (по каждому из концов, причем субъединицей, связанной с противоположным концом, то есть в *транс*-манере). В той же *транс*-манере транспозаза вводит полученные свободные концы в сайт-мишень

АТФ. Присутствие MuB и АТФ строго направляет транспозицию в плазмиду-мишень.

Продукт всех упомянутых реакций — перекрестная структура, в составе которой транспозон связан с мишенью одной цепью с каждого конца. Следующий этап реакции дифференцирован по типу транспозиции. В следующих двух разделах мы увидим, как одна и та же структура может послужить субстратом для репликации (что будет приводить к репликативной транспозиции) или непосредственно использоваться для разрыва и воссоединения в соответствии с нерепликативным механизмом перемещения элемента.

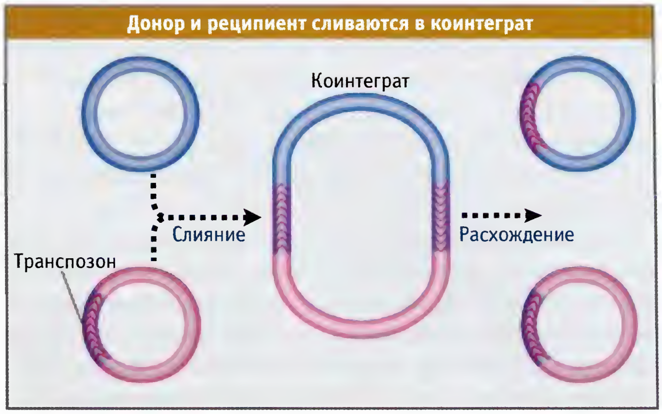


РИС. 21.13 При транспозиции репликоны донора и реципиента могут слиться и образовать коинтеграт. Разделение коинтеграта освобождает репликоны друг от друга, причем каждый из них получает свою копию транспозона

21.7 Репликативная транспозиция сопровождается образованием коинтеграта

Ключевые понятия:

- Репликация в сочетании с переносом цепей порождает коинтеграт — продукт слияния репликонов донора и мишени.
- Коинтеграт содержит две прямых копии транспозона, разделяющих собой последовательности исходных репликонов.
- Рекомбинация между копиями транспозона разделяет коинтеграт на исходные репликоны, причем реципиент получает копию транспозона.
- Реакцию рекомбинации катализирует резольваза, кодируемая транспозоном.

ДНК-структуры, на которых базируется процесс репликативной транспозиции, и их взаимные превращения схематизирует РИС. 21.13:

- После того как транспозаза, согласно описанию из предыдущего раздела, ковалентно свяжет транспозон с выступающими одноцепочечными 5′-концами ДНК-мишени, 3′-концы ДНК-мишени начинают работать затравками (праймерами) для репликации. Двусторонняя репликация ДНК транспозона в кроссовере приводит к образованию **коинтеграта** — двухцепочечного продукта слияния исходных молекул. В его составе — две копии транспозона, разделяющие материал донора и мишени и ориентированные как прямые повторы. Преобразование кроссовера в коинтеграт обеспечивается репликативными системами клетки-хозяина.
- Гомологичная рекомбинация между двумя копиями транспозона высвобождает два обособленных репликона, каждому из которых достается по одной копии транспозона. Один из этих репликонов — донор в исходном виде. Другой — репликон-мишень, заполучивший копию транспозона, окаймленную короткими прямыми повторами сайта внедрения. Реакция рекомбинации называется **расхождением**, за нее отвечает фермент **резольваза**.

У фага Mu образование коинтеграта подробно изучено. Его иллюстрирует РИС. 21.14. Репликоны до-



РИС. 21.14 Транспозиция фага Mu идет через образование кроссовера; репликация превращает его в коинтеграт

нора и мишени удерживаются вместе по дуплексному участку транспозона. Судьба этой структуры соответствует способу транспозиции. Принцип репликативной транспозиции состоит в том, что синхронная двусторонняя репликация дуплицирует транспозон (и, пройдя его, останавливается). Продукт представляет собой коинтеграт.

Перекрестная (кроссоверная) структура, образованная в результате реакции переноса цепей, содержит по обе стороны от транспозона одноцепочечные участки, соответствующие пришитым к транспозону концевым выступам смещенного двухцепочечного разреза ДНК мишени. Эти участки — псевдорепликативные вилки с открытым участком матрицы для синтеза ДНК, который затравляется комплементарными цепями мишени со свободным для удлинения 3′-ОН концом.

Репликация, начавшись в обеих псевдорепликативных вилках, пройдет через транспозон, разделяя его цепи, а затем остановится, дойдя до дуплекса. Получится коинтеграт с прямыми копиями транспозона, соединяющими два репликона (в чем можно убедиться, обойдя контур на рисунке).

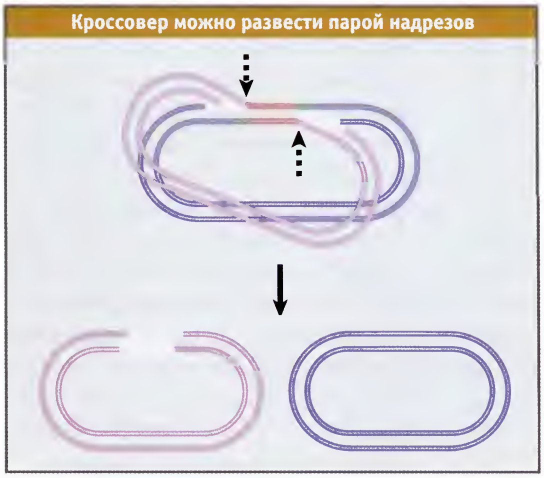


РИС. 21.15 Если кроссовер развести надрезами, транспозиция получится нерепликативной. Транспозон окажется вставленным в сайт-мишень и фланкированным прямыми повторами его нуклеотидной последовательности. Донор останется с двухцепочечным разрывом на месте изъятого транспозона

21.8 Нерепликативная транспозиция сопровождается разрывом и воссоединением

Ключевые понятия:

- **Нерепликативная транспозиция сопровождается разрезанием перекрестной структуры по двум интактным цепям донора и лигированием образовавшихся концов транспозона с оставшимися концами мишени.**
- **При нерепликативной транспозиции Mu первая пара разрезанных цепей оказывается пришита к мишени до того, как разрезана вторая; у Tn5 и Tn10 все четыре разреза вносятся до пришивания к мишени какого-либо из концов.**

Кроссоверная структура используется для нерепликативной транспозиции. В данном случае принцип состоит в разрыве двух оставшихся связок между транспозоном и донором и воссоединении транспозона с мишенью; донор остается разорванным, а коинтеграт не образуется.

РИС. 21.15 изображает нерепликативную транспозицию фага Mu. Как только транспозон будет отрезан от донора по двум оставшимся цепям, его освободившиеся концы могут быть пришиты к свободным концам мишени. Одноцепочечные бреши — следствие смещенности инициирующего разреза мишени — должны быть заполнены репаративным синтезом ДНК. Итоговый продукт — репликон-мишень с транспозоном, вставленным между повторами последовательности, заключавшейся между ини-



РИС. 21.16 В ДНК Tn10 вносятся (с небольшой разницей во времени) две пары надрезов; освободившийся транспозон встраивается в подготовленный сайт-мишень

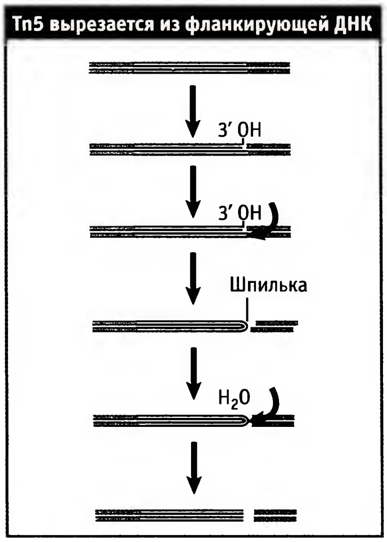


РИС. 21.17 Вырезание Tn5 из фланкирующей ДНК состоит из реакций внесения надреза, межцепочечной атаки и избавления от шпильки

циирующими надрезами мишени. Репликон-донор остается разорванным по обеим цепям в том месте, где раньше была вставка.

При альтернативном варианте нерепликативной транспозиции, параллельно внесению в мишень смещенных надрезов, транспозон полностью высвобождается из окружения донора посредством двусторонних двухцепочечных разрывов (как на рис. 21.7). Путь «надрез и вставка» («cut and paste») свойственен транспозону Tn10, это изображено на **РИС. 21.16**.

Нерепликативный характер транспозиции Tn10 доказывается при помощи красивого эксперимента, использующего искусственно сконструированный дуплекс Tn10 с внесенными в него точечными некомплементарностями. Если транспозиция репликативна, то после нее транспозон придет в строгое соответствие одной из цепей исходного (матричного) Tn10. Если же транспозиция представляет собой физический перенос материала, то все некомплементарности сохранятся. Как было установлено, имеет место второй из этих вариантов.

Главное отличие рис. 21.16 от модели на рис. 21.15 в том, что обе цепи Tn10 разрезаются до какого бы то ни было контакта с сайтом-мишенью. Первым этапом распознавания концов транспозона транспозазой является формирование белковой структуры, внутри которой произойдет реакция. По обеим границам транспозона цепи разрезаются в определенном порядке: первой разрезается перемещаемая цепь (которая в дальнейшем присоединится к мишени), за ней следует оставшаяся цепь. (Этот порядок аналогичен порядку транспозиции у фага Mu, см. рис. 21.14 и рис. 21.15.)

Транспозиция Tn5 также происходит по нерепликативному механизму. На **РИС. 21.17** показана очень любопытная реакция, отделяющая Tn5 от донорских флангов. Надрезание по одной из цепей высвобождает 3'-OH конец, который атакует вторую цепь ДНК. Тогда фланкирующий дуплекс отделяется, а транспозон на конце замыкается в виде шпильки. Верхушку шпильки атакует активированная молекула воды и вскрывает ее. Свободные концы цепей транспозона присоединяются к соответствующим концам мишени. При этом и транспозон, и мишень пребывают заключенными в протеиновой структуре, образованной транспозазой вместе с другими белками. Двухцепочечные разрезы по границам транспозона с донорскими флангами пресекают любой вид репликативной транспозиции, направляя процесс в русло нерепликативной транспозиции. Исход тот же, что и у процесса на рис. 21.14, но обусловлен он другой последовательностью разрезаний и склеиваний отдельно взятых цепей ДНК.

Каждая из транспозаз Tn5 и Tn10 функционирует в форме димера. Каждая субъединица такого димера имеет активный центр, который последовательно катализирует разрывы двух цепей по одной из границ транспозона, а затем участвует в смещенном разрезании мишени. На **РИС. 21.18** показано строение транспозазы Tn5 в комплексе с транспозоном, вырезанным ею из донора. Каждая оконечность транспозона находится в активном центре одной из субъединиц. Другим своим доменом субъединица контактирует с противоположным краем транспозона, что служит пространственной координации процесса. Каждый из активных центров разрезает одну цепь мишени, причем расстояние (смещение) между надрезами задается геометрией комплекса и для Tn5 равняется 5 пн.



РИС. 21.18 Каждая субъединица транспозазы Tn5 удерживает в своем активном центре один из концов транспозона, в то же время контактируя с противоположным краем транспозона

21.9 В транспозиции TnA участвуют транспозаза и резольваза

Ключевые понятия:

- Репликативная транспозиция TnA требует участия транспозазы (для образования коинтеграта) и резольвазы (для разделения репликонов донора и мишени).
- Резольваза по своему действию напоминает белок Int фага лямбда; по своему типу реакция относится к топоизомеразоподобным сайт-специфичным рекомбинациям с образованием промежуточной ковалентной сшивки между белком и ДНК.

Для членов семейства TnA репликативная транспозиция — единственный способ обеспечения их мобильности. Семейство состоит из крупных (>5 тпн) транспозонов, и они не используют, подобно композитным транспозонам, IS-модули, но имеют собственные гены для обеспечения всех своих функций: и транспозиции, и других (например, устойчивости к антибиотикам). Члены семейства обладают типичными структурами из инвертированных близкородственных повторов длиной ~38 пн. Делеция хотя бы по одному из повторов может препятствовать транспозиции. Эти делеции *цис*-действующие (то есть не препятствуют перемещениям других копий транспозона в том же геноме). Интеграция членов семейства порождает прямые повторы сайта-мишени в 5 пн. Характерные для них генетические маркеры резистентности обусловливают сопротивляемость клетки антибиотикам (например, ген *amp*^{r} придает устойчивость к ампициллину). Из членов семейства TnA наиболее полно охарактеризованы Tn3 и Tn1000 (также известный как γδ).

Транспозиция Tn3, опосредованная TnA, проходит в две стадии, за которые отвечают транспозаза и резольваза. Их гены, *tnpA* и *tnpR*, отмечены рецессивными (то есть *цис*-действующими, не препятствующими перемещениям других копий транспозона в том же геноме) мутациями. В собственно транспозиции, как и у IS-содержащих транспозонов, задействованы края и концы элемента. Разделение требует участия специализированного внутреннего сайта, что является уникальной особенностью семейства TnA.

Элементы, мутантные по *tnpA*, не способны перемещаться по геному. Продукт *tnpA* — транспозаза, которая связывается с последовательностью в ~25 пн, лежащей внутри 38 пн инвертированного концевого повтора. Вплотную к ней расположен сайт связывания белка IHF *E. coli*, так что транспозаза с IHF связывают ДНК кооперативно. Транспозаза опознает края элемента, а также вносит в ДНК-мишень надрезы со смещением в 5 пн. IHF — это ДНК-связывающий белок, который часто участвует в монтаже крупных протеиновых агрегатов у *E. coli*; его участие в реакции транспозиции, возможно, не обязательно.

Функция продукта гена *tnpR* двойственна. Он действует и как репрессор транскрипции, и как резольваза.

Мутации в *tnpR* увеличивают частоту транспозиций по той причине, что в норме TnpR репрессирует транскрипцию *tnpA* (равно как и своего собственного гена). Когда TnpR инактивирован, избыточный синтез TnpA приводит к повышенной частоте транспозиций. Это означает, что концентрация TnpA для транспозиции — лимитирующий фактор.

Гены *tnpA* и *tnpR* экспрессируются дивергентно по отношению к межцистронной регуляторной A-T-богатой области, обозначенной на карте рис. 21.19. Все функции TnpR обусловлены его связыванием с ДНК в этой области.

В качестве резольвазы TnpR участвует в рекомбинации между прямыми копиями Tn3 в составе коинтеграта. В принципе развести коинтеграт можно было бы гомологичной рекомбинацией с междуплексным переходом цепей в любой точке транспозона. Однако реакция задействует специализированный сайт, *res*. Он промаркирован делециями, блокирующими завершение транспозиции, что выражается в накоплении коинтегратов. В отсутствие сайта *res* реакцию расхождения может в некоторой степени заменить RecA-опосредованная гомологичная рекомбинация, которая в данном случае гораздо менее эффективна.

Сайты ДНК, которые связывает TnpR-резольваза, представлены в нижней части рис. 21.19. Каждый из них, протяженностью 30–40 пн, связывается с TnpR независимо от остальных двух. Последовательности всех трех сайтов гомологичны и вписываются в симметричный консенсус.

Сайт I включает регион, генетически определенный как *res*-сайт. В его отсутствие специализированная реакция разводки вообще не происходит. Впрочем, расхождение задействует также сайты II и III, поскольку при делеции любого из них реакция идет плохо. Сайт I содержит старт транскрипции *tnpA*, сайт II — старт транскрипции *tnpR*. Мутация, аннулирующая сайт II как сайт эффективной посадки репрессора, картируется на его левом краю.

Взаимодействуют ли сайты I, II и III между собой? Взаимодействие может заключаться в том, что связывание TnpR по всем трем сайтам требуется для придания ДНК про-

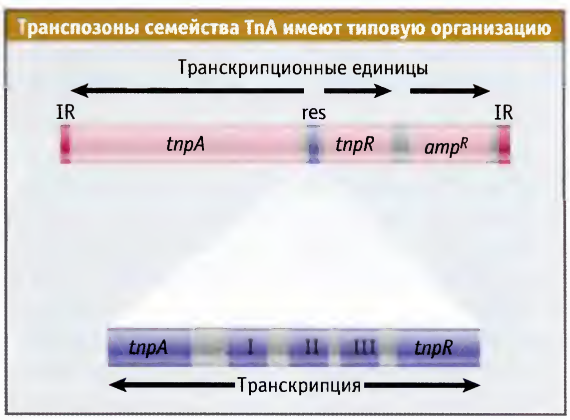


РИС. 21.19 Транспозоны семейства TnA содержат инвертированные концевые повторы (IR), внутренний *res*-сайт, а также три хорошо изученных гена

странственной конформации, необходимой для расхождения. Связывание по отдельным сайтам может репрессировать транскрипцию *tnpA* и *tnpR*, не меняя топологии ДНК.

В экспериментальной системе для изучения разводки *in vitro* субстратом служит похожая на коинтеграт молекула ДНК. Субстрат должен быть сверхспирализован; его разводка порождает два сцепленных кольца, по одному сайту *res* в каждом. Реакция требует больших количеств TnpR-резольвазы и никаких факторов хозяина. Весь процесс происходит в обширном нуклеопротеиновом комплексе. Сначала субъединицы резольвазы связывают сайты *res* в каждой из копий транспозона, а затем связанные сайты сводятся вместе для образования структуры диаметром ~10 нм. Иногда эту структуру называют синаптосомой, и состоит она из шести димеров резольвазы и двух сайтов *res*. Реакция сопровождается изменениями сверхспирализации.

Технически, расхождение представляет собой разрыв и воссоединение связей без дополнительных энергозатрат. Разрыв выполняется по методу трансэтерификации, которая ковалентно пришивает Ser^{10} резольвазной субъединицы к 5′-фосфату разрываемой межнуклеотидной связи в сайте *res*. Промежуточный продукт состоит из резольвазы, ковалентно пришитой к обоим 5′-концам двухцепочечного разрыва, который симметрично делит короткий палиндром 5′TTATAA3′ с образованием динуклеотидных 3′-концевых выступов:

```
        5′TTATAA3′
        3′AATATT5
             ↓
5′TTAT      + белок-AA3′
3′AA-белок       TATT5′
```

Расшифровка нескольких вариантов кристаллической структуры резольвазы γδ показала, что изолированный фермент может образовывать димер димеров в том смысле, что у каждой резольвазной субъединицы есть поверхность, проявляющая сродство к другой субъединице, — так образуется димер, который, в свою очередь, использует вторую пограничную поверхность, чтобы вместе с другим димером образовать тетрамер. Димер димеров может связаться с сайтом *res* так, что аминокислотные остатки активного каталитического центра окажутся пространственно разнесены с тем местом ДНК, которое им предстоит разрезать. Кристаллическая структура синаптосомы, в которой тетрамер резольвазы через остатки Ser^{10} сшит с двумя разрезанными сайтами I ДНК-дуплексов, свидетельствует о первоначальном местонахождении сайтов рекомбинации на противоположных сторонах тетрамера. Это означает, что один из димеров должен повернуться на 180° по отношению к другому димеру — только тогда соответствие положений молекул ДНК сделает рекомбинацию возможной. Эта схема отличается от схемы действия тирозиновых рекомбиназ вроде Cre или Flp, которые соединяют рекомбинирующие ДНК в структуру Холлидея (см. разд. 19.18 «Сайт-специфичная рекомбинация напоминает топоизомеразную активность»).

Резольвазная реакция TnA напоминает действие Int фага лямбда в сайтах *att* (см. разд. 19.16 «Специализированная рекомбинация задействует специфичные сайты»). В самом деле, в сайте *res* есть 20 пн, 15 из которых идентичны основаниям на соответствующих позициях в *att*. Должно быть, сайт-специфичная рекомбинация лямбда и расхождение TnA — эволюционные производные одной и той же реакции рекомбинации. В разд. 23.10 «Белки RAG катализируют разрыв и воссоединение ДНК», мы увидим, что рекомбинация иммуноглобулиновых генов имеет тот же структурный базис. Для всех трех реакций характерен перенос свободного конца ДНК на каталитический белок — как промежуточный этап переподсоединения к другому свободному концу (см. разд. 19.18 «Сайт-специфичная рекомбинация напоминает топоизомеразную активность»).

Сами реакции аналогичны по манипуляциям с ДНК, хотя в расхождении участвуют сайты одной молекулы, в то время как рекомбинация по *att* — межмолекулярная и направленная (что следует из различия сайтов *attB* и *attP*). Механизм действия ферментов тоже отличается. У резольвазы четыре субъединицы связываются с рекомбинирующими сайтами *res*, и каждая субъединица надрезает единственную цепь ДНК. Затем реорганизация субъединиц друг относительно друга физически двигает цепи ДНК, переводя их в рекомбинированную конфигурацию. Процесс завершается заклеиванием стыков и высвобождением резольвазы.

21.10 Контроль транспозиции Tn10 осуществляется различными путями

Ключевые понятия:

- Мультикопийное ингибирование снижает вероятность встраивания каждой новой копии транспозона, если другие копии того же транспозона уже встроились в геном.
- Различные молекулярные механизмы влияют на частоту транспозиций.

Контроль частоты транспозиции важен для клетки. С одной стороны, чтобы вообще подлежать рассмотрению в этой главе, транспозон должен перемещаться с некоторой минимальной частотой. С другой стороны, слишком интенсивная транспозиционная активность может стать губительной для клетки-хозяина. По всей видимости, каждый транспозон обладает каким-либо механизмом контроля частоты собственных транспозиций. Рзнообразие таких механизмов можно рассмотреть на примере Tn10.

Tn10 — это композитный транспозон, перемещение которого обеспечивает его активный модуль IS10R. На РИС. 21.20 показано, как организован этот модуль. Поблизости от внешней границы в нем находятся два промотора. Один из них, P_{IN}, отвечает за транскрипцию

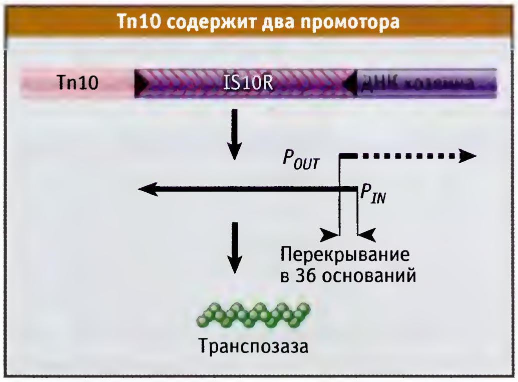


РИС. 21.20 У внешней границы IS10R находятся два противоположно ориентированных промотора. Сильный промотор P_{OUT} поддерживает транскрипцию в направлении фланкирующей ДНК (хозяина). От более слабого промотора (P_{IN}) зависит синтез другой РНК, идущей за пределы IS10R во внутреннюю область транспозона; при ее трансляции получается транспозаза

IS10R. Другой промотор, P_{OUT}, направляет транскрипцию в противоположную сторону — к близлежащей границе транспозона. Обычно транскрипция с P_{OUT} успевает терминироваться в пределах транспозона, но время от времени уходит и дальше, в хромосомное окружение; иногда продолжающаяся транскрипция становится причиной активации смежных с транспозоном бактериальных генов.

Феномен «мультикопийного ингибирования» свидетельствует, что экспрессия транспозазы IS10R подлежит регуляции. Введение в бактериальную хромосому при помощи плазмиды-донора дополнительных копий IS10R приводит к спаду транспозиционной активности Tn10. Ингибирование происходит на уровне трансляции, для него необходим промотор P_{OUT}. Эффект базируется на перекрывании 5′-концевых участков у P_{IN}- и P_{OUT}-транскриптов. РНК OUT — транскрипт со стандартной длиной в 69 оснований. Ее концентрация в 100 раз больше концентрации РНК IN в силу двух обстоятельств: во-первых, P_{OUT}, как промотор, гораздо сильнее, чем P_{IN}; во-вторых, РНК OUT стабильнее, чем РНК IN.

РНК OUT функционирует как антисмысловая РНК (см. разд. 13.7 «Регуляция трансляции с помощью малых молекул РНК»). Если копия транспозона в геноме единственная, то РНК OUT неэффективна; значительный эффект проявляется, когда число копий транспозона достигает пяти. На каждую копию IS10 присутствует ~5 копий РНК OUT (что при обычной степени размноженности Tn10 в геноме соответствует примерно 150 копиям РНК OUT на клетку). РНК OUT и РНК IN взаимодействуют комплементарными 5′-концевыми участками; избыток первой гарантирует быстрое связывание второй — до того как к ней успеет прикрепиться рибосома. РНК IN, взаимодействующая с РНК OUT, не может быть транслирована.

Важнейшим фактором успеха транспозиции является доступность достаточного количества транспозазы. Композитный элемент Tn10, чья транспозаза синтезируется на низком уровне (в 0,15 молекулы на клетку на поколение), обнаруживает несколько интересных способов снижения ее экспрессии, чтобы ограничить оккупацию генома собственными копиями. Эти способы резюмирует РИС. 21.21.

Представленность транспозазы в клетке контролирует скорость транспозиции, мутации по гену транспозазы (с непрерывной рамкой считывания в IS10R) с трудом могут быть компенсированы *транс*-активностью оставшегося IS10R — элемента дикого типа. Это отражает строгое предпочтение транспозазы к *цис*-действию — фермент функционирует эффективно только с той ДНК, которая послужила матрицей для его транскрипции и трансляции. *Цис*-предпочтение вообще характерно для транспозаз, кодируемых IS-элементами (а также для белка А, участвующего в репликации фага φX174, см. разд. 16.5 «Геномы некоторых фагов реплицируются по типу «катящегося кольца»).

Отражает ли *цис*-предпочтение транспозазы ее большее сродство к последовательностям-мишеням, лежащим недалеко от места синтеза фермента? Одно из возможных объяснений состоит в том, что после (или во время) своего синтеза на рибосоме транспозаза привязывается к ДНК так крепко, что вероятность куда-либо диффундировать для нее крайне низка. Другим возможным объяснением служит крайняя нестабильность фермента, не связанного с ДНК, поэтому молекулы белка, которые не успели сразу после синтеза присоединиться к матрице, уже никогда не проявят свою активность.

Благодаря мультикопийному ингибированию и *цис*-предпочтению транспозазы рост числа копий

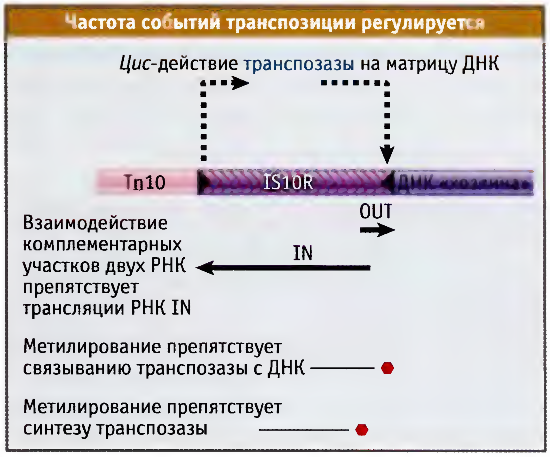


РИС. 21.21 Частота транспозиций Tn10 ограничена сразу несколькими способами. Каждый из механизмов по-своему затрагивает синтез либо функцию белка транспозазы. Для переноса копии транспозона необходимо, чтобы ДНК не была полностью метилирована, что имеет место только после репликации. Выбор мишени ограничивается *цис*-предпочтениями транспозазы, а спаривание комплементарных участков РНК OUT/IN ингибирует синтез транспозазы

Tn10 в бактериальном геноме рано или поздно замедляется и не приводит к увеличению частоты транспозиций, губительной для хозяйского генома.

Для индивидуального элемента (безотносительно остальных копий) самую важную систему регуляции обеспечивают эффекты метилирования. Они не только уменьшают частоту транспозиций, но и (что более важно) приурочивают транспозицию к прохождению репликативной вилки.

Способность IS10 к транспозиции соотнесена с репликационным циклом бактериальной хромосомы ответом транспозона на статус метилирования ДНК по двум сайтам. Один сайт находится внутри инвертированного концевого повтора IS10R, с которым связывается транспозаза. Другой сайт метилирования находится в промоторе P_{IN}, с которого транскрибируется ген транспозазы.

Оба этих сайта метилирует система *dam*, описанная в разд. 15.5 «После репликации ориджины могут быть изолированы от аппарата обеспечения их функции». Метилаза Dam модифицирует аденин последовательности GATC, новосинтезированной в процессе репликации. В штаммах dam^- частота транспозиций Tn10 в 1000 раз выше. Это обусловлено тем, что оба сайта лишены метильных групп.

Прохождение репликативной вилки через эти (метилированные по обеим цепям) сайты порождает полуметилированные сайты дочерних дуплексов; в этом состоянии транспозон активируется, во-первых, за счет учащения актов транскрипции гена транспозазы и, во-вторых, за счет усиления связывания транспозазы с краем IS10R. В штаммах дикого типа сайты остаются полуметилированными лишь на короткий период после репликации. Почему этот период так оптимален для транспозиции?

Нерепликативный механизм транспозиции Tn10 представляет собой угрозу разрушения для ДНК-донора (см. рис. 21.7). Смягчить эту угрозу с точки зрения выживания клетки может вторая копия донора — продукт только что произошедшей репликации. Такое смягчение эффективно, поскольку к транспозиции стимулируется лишь одна из новореплицированных копий Tn10 (у которой конкретная цепь не метилирована по сайтам *dam*).

Транспозон выбирает сайт-мишень случайным образом и с разумной вероятностью может встроиться в активный оперон. Может ли транскрипция извне пройти сквозь транспозон, активировав таким образом транспозазу, перепроизводство которой может, в свою очередь, привести к повышенной (возможно, летальной) частоте транспозиций? У Tn10 существует два механизма защиты от подобных событий. Проход транскрипции через концевую часть снижает его мобильность (предположительно, ингибируя связывание с транспозазой), а мРНК, начинающаяся не с промотора, а выше, транслируется плохо, поскольку имеет такую вторичную структуру, что инициирующий кодон оказывается вне доступа.

21.11 Контролирующие элементы вызывают разломы и перестройки в геноме кукурузы

Ключевые понятия:

- Явление транспозиции было открыто у кукурузы благодаря разломам хромосом, образующимся при транспозиции мобильных «контролирующих элементов».
- Разлом порождает хромосому с центромерой и ломаным концом, а также ацентрический фрагмент (обломок).
- Ацентрический фрагмент теряется во время митоза, это можно заметить по исчезновению доминантных аллелей в гетерозиготе.
- Слияние между ломаными концами двух хромосом порождает дицентрическую хромосому, вовлекаемую в последующие циклы разломов и слияний.
- Цикл разлома-слияния-моста приводит к проявлениям соматической мозаичности.

Самым колоритным образом транспозоны проявляют себя в соматических вариациях, возникающих в онтогенетическом развитии цветковых растений из-за изменений в расположении или поведении **контролирующих элементов**. Такое название было дано транспозонам до того, как была открыта их молекулярная природа.

Кукуруза обладает двумя особенностями, позволяющими отслеживать явления транспозиции. Во-первых, контролирующие элементы часто внедряются поблизости от генов, которые в результате интеграции инактивируются с нелетальным, но фенотипически ярким эффектом. Во-вторых, кукурузе свойственна клональная схема развития, вследствие чего можно отследить события транспозиции во времени и пространстве, как это показывает диаграмма на РИС. 21.22.

Предположим, в одной из клеток зачатка органа (например, зерна) произошла мутация или геномная перестройка. Для визуализации изменения не имеет значения его природа (а это может быть точечная мутация, вставка, делеция или разлом хромосомы), важно только, чтобы событие перестройки меняло экспрессию одного из аллелей в локусе, по которому развивающийся организм гетерозиготен. В этом случае потомство клетки-хозяйки события, став носителем нового фенотипа, будет визуально отличаться от потомства других клеток с исходным фенотипом.

Серия митозов клетки, подвергшейся генетической перестройке, даст начало видоизмененному **сектору** ткани. Изменение фенотипа в процессе соматического развития называют **мозаичностью**; мозаичность проявляется в виде сектора с новым фенотипом, пребывающего в окружении ткани с исходным фенотипом. Размер сектора зависит от числа породивших его клеточных делений, следовательно, соотношение экспансий нового и старого фенотипов (например, на поверхности початка или отдельного зерна) диктуется временем совершения перестройки (хронометражом). Чем раньше в индивидуальном развитии (чем ближе к основанию дерева клеточных делений) произошла мутация, тем больше число несущих ее клеточных потом-

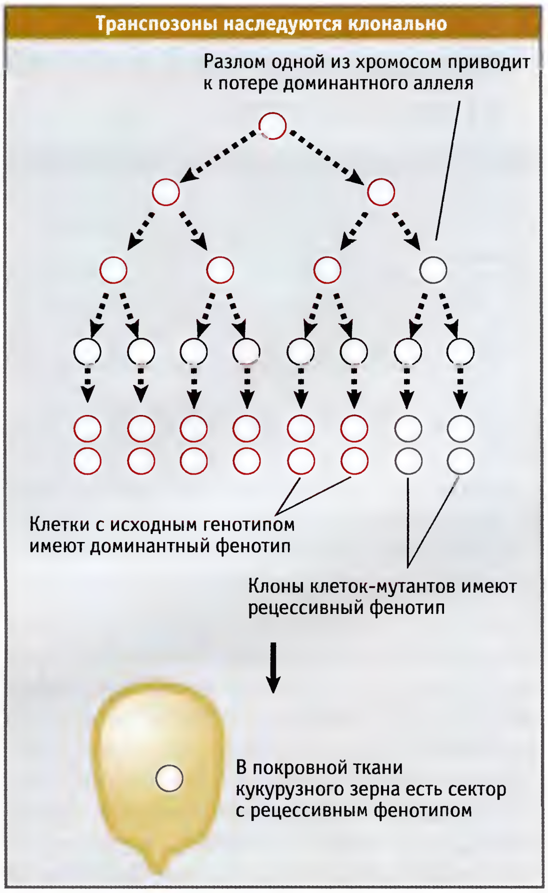


РИС. 21.22 Анализ ситуации, изображенной на рисунке, сводится к идентификации клеточного клона (потомков одной клетки), фенотип которого является результатом события транспозиции. О положении этого события во времени развития зачатка органа можно судить по числу клеток с измененным фенотипом. О тканевой специфичности события (точнее, степени дифференцировки клетки-предшественника на момент события) можно судить по локализации клеток этого клона

ков и тем больше размер участка ткани с новым фенотипом, который они образуют. Самый зрелищный тому пример — цветовые вариации зерен кукурузы.

Классическая причина соматической геномной перестройки — вставка контролирующего элемента, затрагивающая активность смежных генов. Делеции, дупликации, инверсии, транслокации — все они обычны в местах присутствия контролирующих элементов. Самым распространенным последствием присутствия ряда элементов является разлом хромосомы. Уникальным свойством кукурузы как модельного биологического объекта является то, что его контролирующие элементы подлежат регуляции сообразно процессу индивидуального развития организма. Мобильные элементы перемещаются по геному и содействуют генетическим перестройкам в соответствии с характеристическими временными окнами и с частотой, предписываемой стадией развития организма.

Типичным в этом смысле образом ведет себя элемент *Ds*, обнаруженный благодаря своей способности индуцировать разломы хромосом. Последствия его активности иллюстрирует РИС. 21.23. Рассмотрим пару гомологичных хромосом. Одна из них несет серию рецессивных маркеров (*C*, *bz* и *wx*). В другую хромосому *Ds* внедрился между серией доминантных генетических маркеров (*CI*, *Bz* и *Wx*) и центромерой. Разлом хромосомы по Ds порождает **ацентрический фрагмент**, уносящий доминантные маркеры. Будучи лишенным центромеры, ацентрический фрагмент теряется клеткой при митозе. Поэтому в дочерних клетках присутствуют лишь три рецессивных маркера — в составе интактного гомолога. Фенотипический эквивалент — изображение кукурузного зерна с рис. 21.22.

РИС. 21.24 показывает, как разлом хромосомы по *Ds* приводит к образованию двух нестандартных U-образных хромосом. Их порождает слияние ломаных концов у продуктов репликации. U-образный ацентрический фрагмент объединяет в себе дистальные части сестринских хроматид (лежащие, по отношению к центромере, за *Ds*; на рисунке они слева). U-образная **дицентрическая хромосома** объединяет в себе остальные части сестринских хроматид с центромерами (на рисунке — справа от *Ds*). Такое строение дицентрической хромосомы обрекает ее на вхождение в классический так называемый цикл **разлома-слияния-моста**.

Проследим за участью дицентрической хромосомы при попытке к сегрегации в митотическом веретене. Две центромеры тянут ее к противоположным полюсам. От

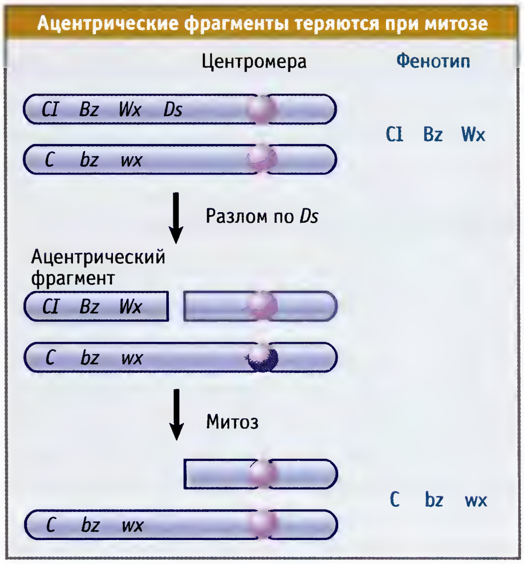


РИС. 21.23 Разлом по контролирующему элементу приводит к потере ацентрического фрагмента хромосомы. Если фрагмент содержит доминантные маркеры, его потеря меняет фенотип гетерозиготы. Фенотипические эффекты потери доминантых маркеров, *CI*, *Bz* и *Wx*, можно оценить по изменению окраски зерен или при помощи специального окрашивания препарированных тканей

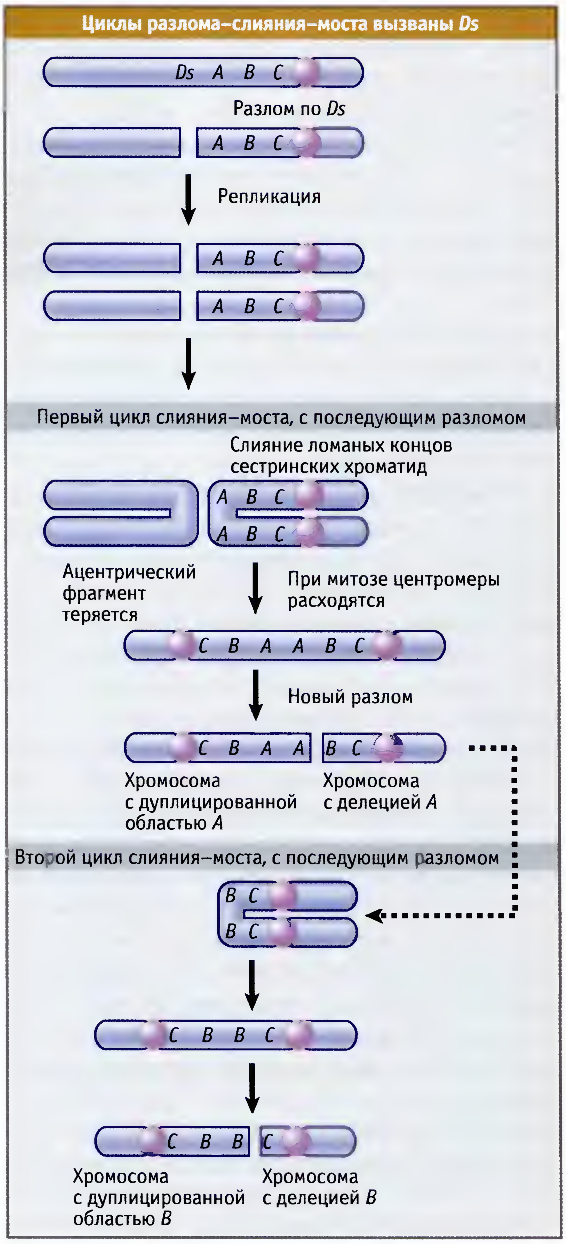


Рис. 21.24 *Ds* служит местом инициации циклических разломов и слияний сестринских хроматид. События молекулярного уровня можно отследить по фенотипам клонов

натяжения хромосома разрывается в произвольном сайте между центромерами. В примере, приведенном на рисунке, разрыв имеет место между локусами *A* и *B* — с тем результатом, что одна из дочерних хромосом получает дупликацию *A*, а другая — его делецию. Если *A* — доминантный маркер, клетки с дупликацией сохранят фенотип *A*, в то время как клетки с делецией перейдут на рецессивный фенотип *a*.

Цикл разлома-слияния-моста повторяется в череде последующих клеточных поколений, генерируя новые генетические изменения. Если развить пример, который только что был рассмотрен, то хромосома с делецией *A* в следующем цикле (при следующем митозе) может претерпеть разрыв между маркерами *B* и *C*, и тогда ее потомки разделятся на тех, что несут дупликацию *B* (на фоне делеции *A*), и тех, у которых *B* (как и *A*) делетирован. Последовательный уход доминантных маркеров (в порядке их расположения на хромосоме по направлению к центромере) фенотипически отображается в виде подсекторов внутри сектора мозаичности.

21.12 Контролирующие элементы образуют семейства транспозонов

Ключевые понятия:

- В каждом семействе транспозонов кукурузы есть автономные и неавтономные контролирующие элементы.
- Автономные контролирующие элементы кодируют белки, позволяющие им перемещаться.
- Неавтономные контролирующие элементы содержат мутации, исключающие наличие у них продуктов, способных к катализу транспозиции; все же они могут перемещаться при условии, что автономный элемент из того же семейства обеспечит их необходимыми белками.
- Автономные контролирующие элементы претерпевают смены фаз функциональности, что обусловлено главным образом динамикой метилированности их ДНК.

Геном кукурузы содержит несколько семейств контролирующих элементов. Каждая сортовая линия растений обладает своим набором элементов с конкретными характеристиками (численностью, типовой принадлежностью, локализацией). Ими может быть оккупирована значительная часть генома. Внутри семейств элементы различаются степенью своей самостоятельности:

- **Автономные контролирующие элементы** способны вырезаться и перемещаться с помощью белков, которые закодированы в них самих. Из-за непрекращающейся активности автономного элемента его внедрение по какому-либо локусу создает нестабильный («мутабельный») аллель. Утрата автономного элемента (или утрата этим элементом своей мобильности) превращает мутабельный аллель в стабильный.
- **Неавтономные контролирующие элементы** сами по себе стабильны (не перескакивают с места на место и вообще не подвержены спонтанным изменениям состояния). Они становятся нестабильны, только если автономный элемент из того же семейства присутствует где-нибудь в другом месте генома. Дополненный *транс*-расположенным автономным элементом неавтономный элемент начинает вести себя активно, демонстрируя весь набор транспозоновых функций, в первую очередь — способность перемещаться в новые сайты. Чаще всего неавтономные элементы — производные автономных элементов, потерявших часть информации, необходимой для воплощения белкового аппарата транспозиции.

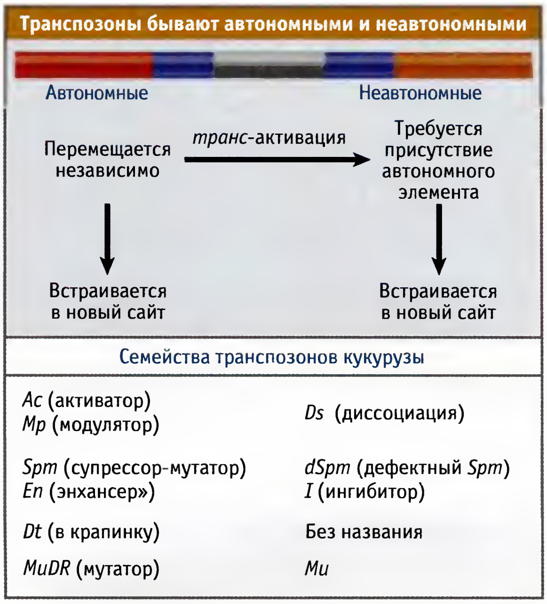


РИС. 21.25 В каждом семействе контролирующих элементов есть автономные и неавтономные члены. Автономные элементы, в отличие от неавтономных, способны к транспозиции вне зависимости от других элементов. Семейств автономных элементов насчитывается не менее четырех; каждое из них дополнено неавтономными элементами

Принадлежность контролирующих элементов к одному семейству определяется возможностью взаимодействия (*транс*-активности) между автономным и неавтономным элементами. Каждое семейство представлено единственной (образующей) разновидностью автономного элемента и многочисленными разновидностями неавтономных элементов. Неавтономный элемент принадлежит семейству, если он может быть активирован образующим автономным элементом из этого семейства. Самые распространенные семейства контролирующих элементов кукурузы перечислены на РИС. 21.25.

Хорошо изученные на молекулярном уровне, транспозоны кукурузы не отличаются какой-то особой формой организации: те же инвертированные концевые повторы, смежные с короткими прямыми повторами флангов. Но по своим размерам и содержанию они сильно варьируют. Для всех семейств характерен один и тот же тип взаимодействия между автономным и неавтономным элементами. У автономных элементов область между концевыми повторами содержит открытые рамки считывания. Неавтономные же элементы никогда не кодируют функциональных белков. Иногда их внутренние области по последовательности похожи на внутреннюю область образующего автономного элемента, в других случаях эти последовательности полностью дивергировали.

Транспозоны семейства *Mutator* — одни из самых примитивных контролирующих элементов. Автономный элемент *MuDR* несет в себе гены *mudrA* (кодирует транспозазу MURA) и *mudrB* (кодирует вспомогательный белок). Края элемента представляют собой инвертированные повторы в 200 пн. Неавтономные элементы семейства — по сути, любая последовательность, окаймленная такими же, как у *MuDR*, повторами, — подлежат мобилизации транспозазой MURA.

Каждое семейство «контролирующих элементов» представлено в геноме растения ~10 членами. Сравнительный молекулярный анализ автономных и неавтономных элементов из семейства *Ac*/*Ds* на основании многочисленных отдельных экземпляров подытожен на РИС. 21.26.

Большую часть автономного элемента *Ac* занимает один-единственный ген. В нем пять экзонов. Его продукт — транспозаза. Элемент окаймляют инвертированные повторы по 11 пн, граничащие с 8 пн дуплицированного сайта внедрения.

Неавтономные элементы *Ds* похожи на *Ac*, но варьируют по последовательности и длине. Их края содержат такие же, как у *Ac* 11 пн, инвертированные повторы. Они короче, чем *Ac*, за счет делеций, размер которых варьирует. Так, у *Ds*9 делетировано только 194 пн, а у *Ds*6 сохранилось 2 тпн (по одной с каждого края). Комплексный двойной элемент *Ds* содержит две последовательности *Ds*6, одна из которых вставлена в другую.

Неавтономные элементы имеют неполную внутреннюю область, но сохраняют инвертированные концевые повторы. Неавтономные элементы происходят от автономных в результате делеций (или других изменений), которые, выводя из строя *транс*-активную (т. е. с действием, распространяющимся и на другие элементы) транспозазу, оставляют нетронутыми сайты, по которым эта транспозаза действует. Делеции, терминирующие автономность, различны: от минимальных (но тем не менее инактивирующих) мутаций в *Ac* до весьма внушительных по своим размерам удаленных или перестроенных участков.

Члены *Ds*1 проявляют другую крайность — они настолько коротки, что сходство их с *Ac* ограничивается

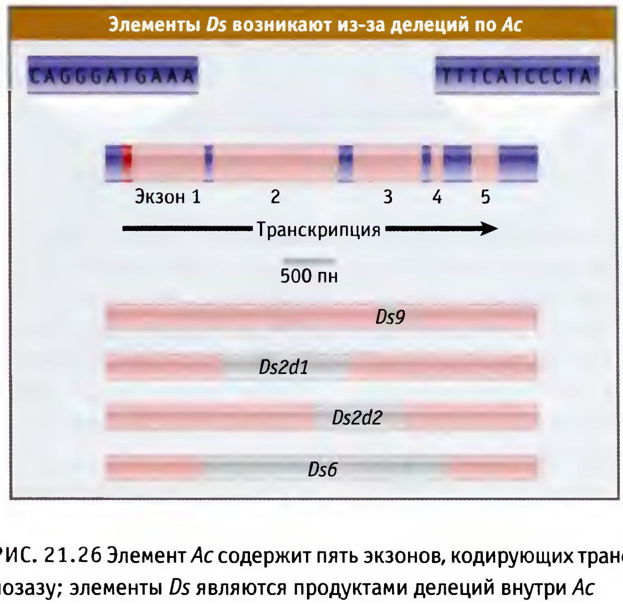


РИС. 21.26 Элемент *Ac* содержит пять экзонов, кодирующих транспозазу; элементы *Ds* являются продуктами делеций внутри *Ac*

только наличием инвертированных концевых повторов. Существование *Ds*1-подобных элементов доказывает, что транспозаза распознает только инвертированные концевые повторы элемента, или, возможно, концевые повторы, связанные с определенными короткими внутренними последовательностями.

По своему механизму транспозиция *Ac/Ds* нерепликативна, то есть сопровождается уходом элемента из донорского окружения. Клональный анализ показывает, что транспозиция *Ac/Ds* почти всегда происходит вскоре после того, как элемент в составе донора оказывается реплицированным. Те же особенности присущи транспозиции бактериального элемента Tn10 (разобранной в разд. 21.10 «Транспозиция Tn10 подвержена разным формам контроля»). Действующая причина в обоих случаях одна и та же: транспозиции не происходит, когда ДНК транспозона метилирована по обеим цепям. Полуметилированная ДНК, которая только что прореплицировалась, способствует активизации транспозона. У *Ac/Ds* сайт-реципиент часто расположен на той же хромосоме, иногда — совсем близко от донора.

Репликация порождает две копии потенциального донора *Ac/Ds*, но (как и у Tn10) транспозицию обычно совершает лишь одна из них. Как складывается дальнейшая судьба покинутого сайта-донора? Перестройки, обнаруживаемые в покинутых сайтах, могут быть квалифицированы как последствия разломов хромосом, изображенных на рис. 21.23.

Состояния контролирующих элементов подвержены разнообразным изменениям. Некоторые из этих изменений генетические, другие — эпигенетические.

Главное изменение — это, конечно, превращение автономного элемента в неавтономный. Помимо этого, неавтономный элемент, обретя какой-нибудь дефект, может стать невосприимчив к мобилизации автономным элементом. Такой неавтономный элемент может перманентно стабилизироваться в геноме, поскольку больше не может быть побужден к транспозиции.

Автономные элементы подвержены сменам фазы — наследуемым, но относительно нестабильным изменениям своих свойств. Это что-то вроде обратимой инактивации, когда контролирующий элемент колеблется между активным и неактивным состояниями в течение индивидуального развития растительного организма.

У автономных элементов *Ac* и *MuDR* смена фазы обусловлена сменой статуса метилирования ДНК. Результаты сравнения активных и неактивных элементов на предмет чувствительности к расщеплению эндонуклеазами рестрикции позволяют утверждать, что в своей неактивной форме элемент метилирован по последовательности-мишени $\frac{\text{CAG}}{\text{GTC}}$. В каждом элементе есть несколько таких сайтов, и пока неясно, какие из них участвуют в контроле активации. Деметилирование концевых повторов *MuDR* усиливает экспрессию транспозазы, что можно расценивать как свидетельство участия гена транспозазы в опосредовании стимулирующего эффекта. О том, как контолируются метилирование и деметилирование, хотелось бы, конечно, знать больше.

Эффект метилирования вообще обычен для транспозонов растений. Лучшая его демонстрация принадлежит *ddm1*⁻, мутанту *Arabidopsis*. При мутировавшем *ddm* гетерохроматин неметилирован. Анализ динамики геномных последовательностей показал, что деметилирование транспозонов *Arabidopsis*, родственных *MuDR* кукурузы, приводит к повреждению генома вследствие резкого увеличения частоты транспозиций. Возможно, метилирование и есть главный рычаг, спасающий геном от разрушения из-за слишком частых транспозиций.

У транспозонов растений могут быть и самоуправляемые механизмы контроля, аналогичные иммунитету бактериальных транспозонов, оберегающих окрестности от собственных копий. Так, рост числа копий *Ac* в отдельно взятом геноме снижает частоту транспозиций. Возможно, элемент *Ac* кодирует репрессор транспозиции; возможно, что этим репрессором работает сама транспозаза — в другой функциональной ипостаси.

21.13 Элементы *Spm* влияют на экспрессию генов

Ключевые понятия:

- Элемент из семейства *Spm* может влиять на экспрессию генов по месту своего внедрения, если его края связывает белок TnpA.
- Элемент *Spm* инактивируется метилированием.

Автономные элементы *Spm* и *En* практически идентичны — различаются менее чем по 10 позициям. **РИС. 21.27** дает представление об их структурах. Инвертированные концевые повторы в 13 пн принципиально важны для транспозиции, о чем свидетельствует транспозиционно-дефектный фенотип краевых делеций. Транспозоны, похожие на *Spm*, найдены и у других растений. Их с *Spm* относят к одному семейству на основании общности организации. У всех у них почти одинаковые инвертированные концевые повторы, а при их транспозиции дуплицируется по 3 пн сайта-мишени. За альтернативное название всего семейства — транспозоны CACTA — было взято сочетание букв, обозначающих нуклеотидный консенсус краев элемента.

С промотора, расположенного по левому краю элемента, транскрибируется последовательность в 8300 пн. Одиннадцать экзонов транскрипта сплайсируются в 2500 оснований мРНК, которая кодирует белок из 621 аминокислоты. Ген называют *tnpA*, а кодируемый им белок связывается с многочисленными копиями 12 пн консенсусных последовательностей, присутствующих по краям элемента. Функция *tnpA* необходима (но, возможно, сама по себе недостаточна) для эксцизии транспозона.

Все неавтономные элементы семейства (обозначаемые как *dSpm*, «дефектный *Spm*») по структуре очень



РИС. 21.27 Автономные элементы семейств *Spm* и *En* содержат по два гена. Ген *tnpA* имеет одиннадцать экзонов, которые транскрибируются и склеиваются в мРНК длиной 2500 оснований. Гену *tnpB* соответствует мРНК длиной 6000 оснований с двумя открытыми рамками считывания

похожи на *Spm*. Они несут делеции, затрагивающие экзоны *tnpA*.

Две дополнительных открытых рамки считывания, ORF1 и ORF2, находятся внутри длинного первого интрона *tnpA*. Они содержатся в альтернативно сплайсированном продукте транскрипции — мРНК в 6000 оснований, присутствующем в количестве, в 100 раз меньшем, чем мРНК *tnpA*. Генетически, ORF1 и ORF2 объединены как *tnpB*. Продуктом *tnpB* может быть белок, который связывается с 13 пн инвертированных концевых повторов и вырезает транспозон из донорского окружения.

У полноценного *Spm* есть производные *Spm-w*, чья транспозиционная активность слабее. Пример *Spm-w*, приведенный на рис. 21.27, несет делецию, элиминирующую вместе взятые ORF1 и ORF2. Это говорит о том, что транспозиция может осуществляться без *tnpB* или с его заменой альтернативной функцией.

Вставки *Spm* могут участвовать в контроле экспрессии генов по месту своего внедрения. Локус-реципиент может очутиться под их контролем — как негативным, так и позитивным. По определению, *Spm*-супрессируемый локус претерпевает ингибирование экспрессии, в то время как *Spm*-зависимый локус может быть экспрессирован только при поддержке *Spm*. Если произошла инсерция *dSpm*, супрессия или зависимость тоже могут возникнуть — как ответ на *транс*-активную функцию, осуществляемую автономным *Spm*.

На чем основываются столь противоположные эффекты, как супрессия и зависимость? Супрессируемый аллель содержит вставку *dSpm* внутри экзона. Каким образом такая структура вообще может быть экспрессирована — даже на низком уровне? Дело в том, что последовательность *dSpm* может быть вырезана из транскрипта в ходе сплайсинга, если будут опознаны ее края. Такой нестандартный сплайсинг может видоизменить окончательную последовательность мРНК, что отразится на свойствах кодируемого ею белка. Похожая способность быть вырезанными из транскрипта была обнаружена у некоторых вставок *Ds*.

Супрессирующую функцию, по которой было дано название всему семейству *Spm*, обеспечивает *tnpA*. Присутствие неавтономного *dSpm* может уменьшить экспрессию гена, но не полностью свести ее на нет. Внедрение же автономного элемента, несущего функциональный ген *tnpA*, может полностью подавить экспрессию гена-мишени. Эффект супрессии обусловлен способностью TnpA опознать края элемента и связаться с ними, заблокировав ход транскрипции.

Аллель, зависимый от *Spm* или *dSpm*, находится рядом с сайтом внедрения элемента. По всей видимости, вставка служит энхансером (усилителем) транскрипции с близлежащего промотора (принадлежащего зависимому гену).

Оба эффекта — супрессия и зависимость — опираются, несмотря на свою противоположность, на одно и то же взаимодействие между (*транс*-активным) продуктом гена *tnpA* автономного элемента и сайтами по краям (возможно, другого, неавтономного) элемента, внедрившегося в конкретном месте. Таким образом, одно и то же взаимодействие между белком и ДНК либо подавляет, либо стимулирует локус-мишень, что обусловлено расположением элемента внутри гена или выше него.

Элементу *Spm* свойственно множество различных состояний. Он может быть в полной мере активен или же вести скрытое (криптическое) существование. Криптический элемент «молчит», т. е. не экспрессирует свои гены и тем самым не участвует в транспозиции — ни своей собственной, ни элементов *dSpm*. Криптический элемент может быть временно активирован и даже надолго переведен в активное состояние через взаимодействие с активным элементом *Spm*. Обратное превращение — инактивация — происходит в результате метилирования последовательности ДНК неподалеку от старта транскрипции *tnpA*. В чем заключается молекулярная природа метилирования *de novo* и каковы возможности активации деметилированием (или предотвращением метилирования), пока неизвестно.

21.14 Роль мобильных элементов в гибридном дисгенезе

Ключевые понятия:

- **Особи линий P *Drosophila melanogaster* являются носителями генетических P-элементов, представляющих собой транспозоны.**
- **В результате скрещивания самца P с самкой M транспозиция P-элементов активируется.**
- **Внедрение P-элементов в новые сайты инактивирует множество генов в геноме потомства от скрещивания самца P с самкой M, приводя к бесплодию.**

Некоторые линии плодовых мушек *D. melanogaster* плохо скрещиваются между собой. У их потомства проявляются признаки вырождения, которые на молекулярном уровне представляют собой серийные мутации, хромосомные аберрации, извращенную мейотическую сегрегацию хромосом. Появление этих признаков, коррелирующих между собой и обуславливающих бесплодие (стерильность) особи, называют **гибридным дисгенезом**.

У *D. melanogaster* описано две модельных системы гибридного дисгенеза. В одной из них мухи разделяются на типы I (inducer, индуцирующий) и R (reactive, реагирующий). Сниженная плодовитость свойственна гибридам самцов I с самками R, но не наоборот. Во второй системе мухи делятся по типам P (от *англ.* paternal contributing — отцовский вклад) и M (от *англ.* maternal contributing — материнский вклад). РИС. 21.28 подчеркивает асимметрию системы: скрещивание между самцом P и самкой M приводит к дисгенезу, а обратное скрещивание — нет.

Дисгенез, как феномен, берет начало в клетках зародышевого пути (герминальных клетках). При скрещивании P-M, потомство F1 (первого поколения) обладает нормальными соматическими тканями. Однако гонады у них не развиваются, а морфологические отклонения в развитии гамет фиксируются, начиная со стадии быстрых клеточных делений герминальной линии.

Любая из хромосом самца P может индуцировать дисгенез у гибрида с самкой M. Строение рекомбинантных хромосом показывает, что несколько областей внутри каждой из P-хромосом способны к подобной индукции. Это говорит о том, что в геноме самцов P существует множество сайтов, порождающих индукцию дисгенеза. Индивидуальные линии мушек P отличаются между собой по расположению этих последовательностей, отсутствующих в хромосомах мушек M.

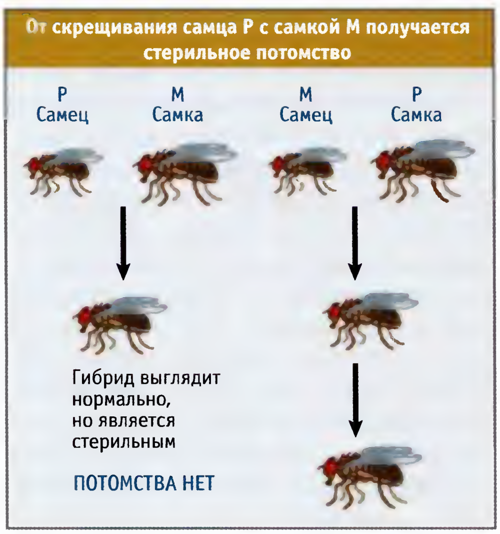


РИС. 21.28 Гибридный дисгенез асимметричен: он проявляется при скрещивании самца P с самкой M, но не самца M с самкой P

Впервые природа P-специфичных последовательностей была установлена при картировании ДНК у мутантов *w*, попадающихся среди дисгенезных гибридов. Все они несут вставку ДНК по локусу *w*. Инактивация локуса вставкой приводит к наличию у мушки белых глаз, почему он, собственно, и назван «белым» (white). А вставленная в него последовательность получила название **P-элемент**.

Вставки P-элемента образуют классическую мобильную систему. Отдельные элементы варьируют по длине, но гомологичны по последовательности. Все они обладают инвертированными концевыми повторами по 31 пн в сопровождении прямых повторов ДНК-мишени по 8 пн. Самые крупные P-элементы имеют длину ~2,9 тпн и содержат четыре открытых рамки считывания. Более короткие элементы возникают, по всей видимости, довольно часто, из-за внутренних делеций полноразмерного P-элемента. По крайней мере, некоторые из укороченных P-элементов потеряли способность продуцировать транспозазу, но могут быть *транс*-активированы ферментом, закодированным в полноценном P-элементе.

Стандартная линия мушек P содержит от 30 до 50 P-элементов на геном, примерно треть их полноразмерны. P-элементы отсутствуют в геноме мушек линий M. Линия P несет P-элементы как инертные компоненты генома, но они активируются, когда самец P скрещивается с самкой M.

Хромосомы гибридных дисгенезных мушек P-M несут P-элементы, внедрившиеся во много новых сайтов. Вставки инактивируют гены, в которые были привнесены, и нередко служат причиной разлома хромосом. Инактивация генома — вот каков общий результат всплеска транспозиционной активности.

21.15 P-элементы могут быть активированы в клетках зародышевого пути

Ключевые понятия:

- При скрещивании самца P и самки M P-элементы в клетках зародышевого пути их потомства активно перемещаются по геному, поскольку при удалении одного интрона в ходе сплайсинга, специфичного для клеток зародышевого пути, образуется последовательность, кодирующая транспозазу.
- Другой белковый продукт P-элемента — репрессор транспозиции, синтезируется во всех клетках и его очень много в цитоплазме яйцеклеток самки P.
- Отсутствие признаков дисгенеза у гибрида между самцом M и самкой P объясняется присутствием репрессора в цитоплазме оплодотворенной яйцеклетки.

Хотя P-элементы транскрибируются и в соматических тканях, их перемещение по геному характерно лишь для клеток зародышевого пути. Тканевая специфичность транспозиции связана с изменением паттерна сплайсинга.

РИС. 21.29 изображает организацию Р-элемента и его транскриптов. Первичный транскрипт в среднем составляет 2,5 тпн, но может доходить и до 3 тпн — вероятно, вследствие сбоев («протекания») терминации. На основе первичного транскрипта может, в конечном счете, быть синтезировано два альтернативных белковых продукта:

- В соматических тканях при сплайсинге вырезаются только первые два интрона, образуя кодирующую область ORF0-ORF1-ORF2. Трансляция получившейся мРНК дает белок массой 66 кДа. Он работает репрессором транспозиционной активности.
- В герминальных тканях дополнительный маневр сплайсинга удаляет интрон 3. Благодаря этому в состав мРНК входят все четыре открытых рамки; транслируемый с нее белок весит 87 кДа. Этот белок — транспозаза.

Необходимость в сплайсинге третьего интрона для транспозиции была доказана экспериментами двух типов. Во-первых, если в сайты сплайсинга (splice junctions, участки на стыках между экзонами и интронами) внести мутации *in vitro*, а затем внедрить эти Р-элементы в геномы мушек, то окажется, что эти элементы потеряли всякую мобильность. Во-вторых, если у Р-элемента третий интрон делетирован, и тем самым ORF3 включена в состав мРНК конститутивно во всех тканях, транспозиция имеет место и в соматических тканях — наравне с герминальными. Таким образом, присоединение ORF3 к предшествующей рамке считывания — необходимое и достаточное условие транспозиционной активности Р-элемента. Это критический аспект регуляции, в норме присущий лишь клеткам зародышевого пути.

Чем обеспечена тканевая специфичность сплайсинга? Соматические клетки содержат белок, который связывается с последовательностью экзона 3, препятствуя сплайсингу (удалению из транскрипта) последнего (третьего) интрона (см. разд. 26.12 «Альтернативный сплайсинг — это реализация различных вариантов сшивания экзонов между собой»). В герминальных клетках такой белок отсутствует, и сплайсинг беспрепятственно генерирует мРНК, кодирующую транспозазу.

Для транспозиции Р-элементу требуются ~150 пн собственной концевой ДНК. Транспозаза связывается с последовательностями в 10 пн, примыкающими (изнутри элемента) к 31 пн инвертированных концевых повторов. Транспозиция происходит по нерепликативному механизму «надрез и вставка» («cut and paste») — как у бактериального Tn10. Такой механизм способствует гибридному дисгенезу двумя способами: вставка в новом месте может задеть интактный локус и вызвать его мутацию, а разорванное состояние покинутого донора способствует возникновению делеций.

Интересно, что в значительном проценте случаев репарация разрыва ДНК донора происходит с привлечением информации (нуклеотидной последовательности) из гомологичной хромосомы. Если гомолог несет в себе Р-элемент, то присутствие Р-элемента в сайте-доноре может быть восстановлено (процессом наподобие репликативной транспозиции). Если у гомолога на этом месте Р-элемент отсутствует, то Р-элемент удаляется из донора, и результатом суперпозиции репарации и транспозиции будет донор, аккуратно зашитый по месту точно вырезанного Р-элемента (при том, что точное вырезание для транспозиции, скорее, редкость).

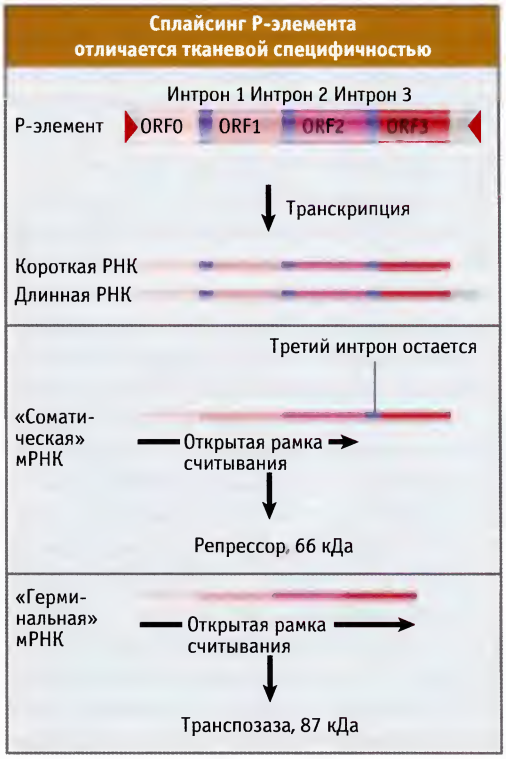


РИС. 21.29 Р-элемент содержит четыре экзона. В герминальных клетках все четыре экзона объединены в одну открытую рамку считывания. В соматических клетках рамка укорочена, поскольку в ней отсутствует материал четвертого экзона

Зависимость гибридного дисгенеза от половой ориентации скрещивания обусловлена еще одной формой наследственности — цитоплазматической. Как известно, вклад самца в оплодотворение ограничивается генетическим материалом, в то время как самка, кроме второй половины генома, представляет также сложнейшим образом структурированную цитоплазму яйцеклетки (ооцита). Вклад цитоплазмы яйцеклетки в наследственность обозначается термином **цитотип**. Линия мушек-носителей Р-элементов имеет цитотип Р, мушки без Р-элементов имеют цитотип М. Гибридный дисгенез случается только вследствие нахождения хромосом с Р-элементами (Р-факторами) в условиях цитотипа М, т. е. только если отцовские Р-элементы попали в яйцо с цитоплазмой М.

Если скрещивание пройдет через цитотип Р (то есть потомство даст самка, у которой в геноме есть Р-элементы), то у самок с генотипом М из ее потомства эффект

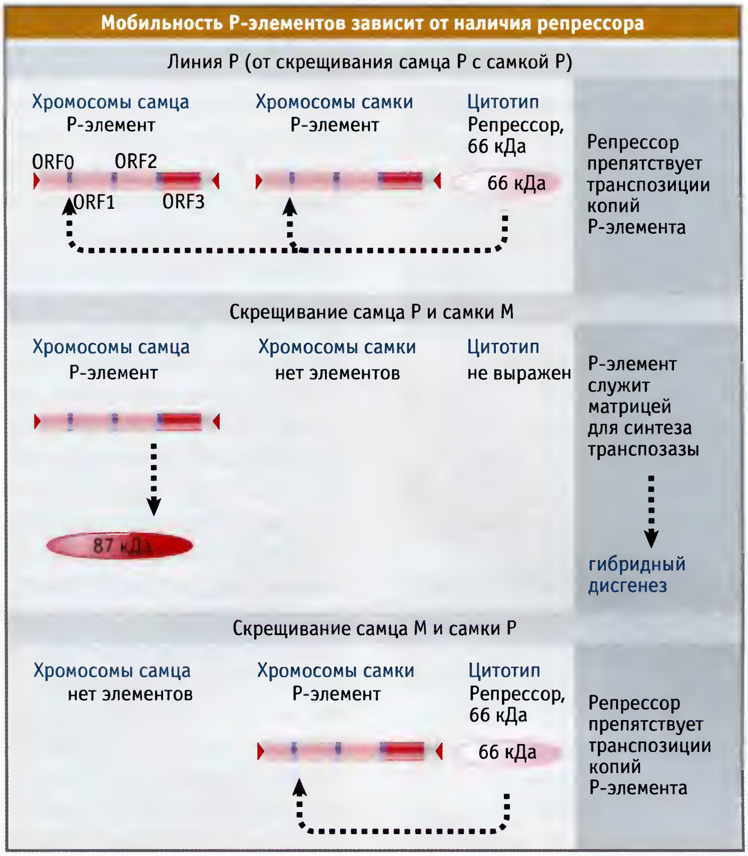


РИС. 21.30 Гибридный дисгенез обусловлен взаимодействием между элементами цитотипа и генома (между репрессором и Р-элементом)

гибридного дисгенеза при скрещивании будет подавлен в нескольких поколениях кряду. Значит, в цитотипе Р есть нечто, подавляющее гибридный дисгенез, что может быть разбавлено цитоплазмой другого типа лишь за несколько поколений.

С молекулярно-биологической точки зрения, эффект цитотипа объясняется моделью на РИС. 21.30. Он определяется способностью белка в 66 кДа репрессировать транспозицию. Этот белок присутствует в яйцеклетках Р как материнский цитоплазматический фактор. У мух линий Р во всех клетках организма этого белка должно быть достаточно, чтобы он препятствовал нежелательной транспозиции. В любом скрещивании с участием самки Р присутствие этого фактора в зиготе (оплодотворенной яйцеклетке) предотвращает синтез и активность транспозазы. Если же яйцеклетка принадлежит самке М и не содержит репрессора, то Р-элемент, внесенный туда как часть отцовской ДНК при оплодотворении, свободен продуцировать транспозазу и пользоваться ею без каких бы то ни было ограничений. Способность цитотипа Р распространять свойственные ему ограничения на несколько поколений вперед говорит о том, что концентрация репрессора в яйце Р должна быть весьма высока. Также этот белок должен быть чрезвычайно стабилен, чтобы дождаться передачи следующему поколению через ооцит развившейся и повзрослевшей особи.

Лабораторные линии *D. melanogaster*, ведущие свое происхождение от диких мушек, пойманных более 30 лет тому назад, всегда М. Если линия происходит от мушек, пойманных в течение последних 10 лет, то она почти всегда является Р-линией. Не означает ли это, что семейство Р-элементов вторглось в геном диких популяций за последние несколько десятков лет? Действительно, Р-элементы очень инвазивны и при попадании в новую для себя популяцию очень быстро ее осваивают; источником инвазивного элемента послужил, вероятно, какой-то другой вид.

Гибридный дисгенез редуцирует межгрупповое скрещивание и потому является шагом на пути к образованию новых таксономических групп. Предположим, в какой-нибудь географической точке земного шара семейство транспозонов образовало дисгеничную систему скрещивания в популяции мушек (или других организмов). Другое семейство элементов может — на основе того же вида мушек — создать другую систему в другой точке. Мушки одного вида, но происходящие с разных территорий, будут дисгеничны по двум системам транспозонов. Создавшаяся таким образом генетическая изоляция готовит почву для дальнейшей сепарации и способствует видообразованию.

21.16 Резюме

Клетки про- и эукариот содержат огромное разнообразие транспозонов, находящихся в движении: их последовательности подлежат перемещению или копированию в альтернативные области генома. По отношению к геному транспозоны всегда находятся внутри, их мобильность не подразумевает независимой формы существования. Транспозон традиционно рассматривается как «эгоистичная ДНК» — элемент генома, озабоченный лишь самоподдержанием; тем не менее во множестве ситуаций транспозоны случайно сообщают хозяину селективное преимущество. Все транспозоны обладают системами ограничения своей мобильности, потому что необузданная транспозиция может причинить клетке значительный ущерб, но молекулярные механизмы такого ограничения могут быть очень разными.

В типичном случае транспозон несет по краям инвертированные повторы и окаймлен прямыми повторами короткой последовательности сайта своего внедрения в геном. Простейшие транспозоны, такие как вставочные последовательности (IS-элементы) бактерий, состоят из инвертированных концевых повторов, фланкирующих (как правило, одну) кодирующую рамку, белковый продукт которой — транспозаза — обеспечивает мобильность элемента. Композитные транспозоны имеют концевые модули, состоящие из IS-элементов; по меньшей мере, один из них способен обеспечить транспозазную активность. С точки зрения транспозиции, область ДНК между модулями (часто она содержит гены устойчивости к антибиотикам) — просто пассажир.

Образование повторов последовательности-мишени, окаймляющих внедрившийся в нее транспозон, отражает характерную черту всех механизмов транспозиции. Сайт-мишень подготавливается к внесению транспозона при помощи двух надрезов (по одному в каждой цепи ДНК), смещенными друг относительно друга на фиксированное расстояние (часто равное 5 или 9 пн). Транспозон вставляется между одноцепочечными выступами краев разреза. Прямые повторы мишени возникают в результате заполнения одноцепочечных участков (брешей).

Транспозиция IS-элементов, композитных транспозонов, Р-элементов подчинена нерепликативному механизму, при котором элемент физически перемещается из сайта-донора в сайт-реципиент (мишень). Реакцию ведет единственный фермент (транспозаза), и делает это в режиме «надрез и вставка» («cut and paste»), при котором транспозон полностью вычленяется из фланкирующей ДНК. Отделение концов транспозона от донора, внесение надрезов в реципиент, подсоединение концов транспозона к смещенным надрезам — все это происходит в составе нуклеопротеинового комплекса, содержащего транспозазу. Уход транспозона оставляет донор в разорванном состоянии. У композитного элемента Tn10 транспозиция привязана по времени к периоду сразу после репликации ДНК. Это обеспечивается благодаря тому, что сайты, находящиеся в ведении системы метилирования *dam*, временно оказываются метилированными только по одной цепи ДНК. Таким образом, транспозиция приходится по времени на момент, когда вся ДНК в клетке продублирована — чтобы не жертвовать обеими копиями донора для транспозиции, которая опасна для целостности клетки.

Семейство транспозонов TnA пользуется репликативным механизмом перемещения. Транспозон, не выходя полностью из состава сайта-донора, подсоединяется к мишени. Взяв за основу эту связку, репликация через область транспозона создает молекулу-коинтеграт, содержащую две копии транспозона. Реакция расхождения (разделения) представляет собой рекомбинацию по конкретному сайту последовательности транспозона, с последующим высвобождением двух его копий: одна остается в доноре, другая оказывается внедренной в сайт-мишень. Весь процесс требует участия двух ферментов, закодированных в транспозоне: транспозаза узнает края мобильного элемента и подсоединяет его концы к сайту-мишени, а резольваза обеспечивает сайт-специфичную рекомбинацию.

Бактериофаг Mu проходит репликативную транспозицию по тому же механизму, что и TnA. Кроме того, он может использовать полупродукт-коинтеграт для нерепликативного перемещения. Отличие реакций нерепликативной транспозиции у фага Mu и IS-элементов заключается в разном порядке разрезания цепей ДНК.

Самые изученные транспозоны растений — это контролирующие элементы кукурузы, которые по систематике распадаются на несколько семейств. В каждом семействе есть единственная разновидность автономного элемента, аналогичного бактериальным транспозонам по своей способности к самостоятельным перемещениям по геному. Кроме него, к семейству относится множество различных неавтономных элементов — производных от автономного, дефектных вследствие мутаций (обычно, делеций). Неавтономный элемент способен совершить транспозицию, только если в той же клетке присутствует автономный элемент из его семейства. Он *транс*-активно предоставляет неавтономному элементу свои функциональные возможности, которыми тот пользуется для совершения транспозиций.

Помимо непосредственных эффектов от интеграции и эксцизии, контролирующие элементы кукурузы могут контролировать активность генов вблизи того места, где находится вставка; такой контроль может подлежать регуляции в соответствии со стадией индивидуального развития растительного организма. Транспозоны кукурузы, вставленные в гены, могут вырезаться из транскриптов — это объясняет, почему, при интеграции в кодирующую рамку, им не удается подавить активность гена. Контроль экспрессии генов кукурузы транспозонами заключается в разнообразных молекулярных эффектах, как то: активация с предоставлением энхансера, супрессия с блокированием хода транскрипции.

Транспозиция мобильных элементов кукурузы (в особенности *Ac*) нерепликативна, и, вероятно, требует участия единственного фермента — транспозазы, кодируемой самим элементом. Транспозиция преимущест-

венно происходит сразу после репликации элемента. По всей видимости, существуют механизмы, ограничивающие частоту транспозиций. Перестройки генома, позитивные с точки зрения кукурузы, могут быть обусловлены присутствием контролирующих элементов.

Мобильные Р-элементы в геноме *D. melanogaster* обусловливают гибридный дисгенез. Подобные явления могут способствовать генетической изоляции между популяциями и, тем самым, вносить серьезный вклад в образование новых видов организмов. Потомство от скрещивания самца, несущего Р-элементы в составе своего генома, с самкой, у которой их нет, стерильно. Р-элемент содержит четыре открытых рамки считывания, разделенных интронами. Сплайсинг, объединяющий первые три рамки, имеет место во всех клетках и приводит к синтезу 66 кДа репрессора. Присутствие в мРНК всех четырех открытых рамок считывания характерно лишь для герминальной клеточной линии, ведущей к образованию половых клеток. При попадании в цитоплазматическое окружение, свойственное яйцеклетке, где нет 66 кДа репрессора, Р-элементы начинают активно перемещаться. Такая мобилизация, носящая взрывной характер, инактивирует геном половых клеток в результате образования многочисленных вставок по случайным положениям. Только полноценный Р-элемент может служить источником транспозазы, которая может также обеспечивать транспозицию дефектных Р-элементов.

Литература

21.1 Введение

Обзор

Campbell, A. (1981). Evolutionary significance of accessory DNA elements in bacteria. *Annu. Rev. Immunol.* 35, 55–83.
Finnegan, D. J. (1985). Transposable elements in eukaryotes. *Int. Rev. Cytol.* 93, 281–326.

21.2 Вставочные последовательности — это простейшие модули транспозиции

Обзор

Berg, D. E. and Howe, M., eds. (1989). *Mobile DNA.* Washington, DC: American Society for Microbiology Press.
Calos, M. and Miller, J. H. (1980). Transposable elements. *Cell* 20, 579–595.
Craig, N. L. (1997). Target site selection in transposition. *Annu. Rev. Biochem.* 66, 437–474.
Galas, D. J. and Chandler, M. (1989). Bacterial insertion sequence. In Berg, D. E. and Howe, M., eds., *Mobile DNA.* Washington, DC: American Society for Microbiology Press, pp. 109–162.
Kleckner, N.(1977). Translocatable elements in prokaryotes. *Cell* 11, 11–23.
Kleckner, N. (1981). Transposable elements in prokaryotes. *Annu. Rev. Genet.* 15, 341–404.

Исследование

Grindley, N. D. (1978). IS1 insertion generates duplication of a 9 bp sequence at its target site. *Cell* 13, 419–426.
Johnsrud, L., Calos, M. P., and Miller, J. H. (1978). The transposon Tn9 generates a 9 bp repeated sequence during integration. *Cell* 15, 1209–1219.

21.3 Композитные транспозоны содержат IS-модули

Обзор

Kleckner, N. (1989). Tn10 transposon. In Berg, D. E. and Howe. M. M., eds., *Mobile DNA.* Washington, DC: American Society for Microbiology Press, pp. 227–268.

21.4 Транспозициям может идти по репликативному или по нерепликативному механизмам

Обзор

Craig, N. L. (1997). Target site selection in transposition. *Annu. Rev. Biochem.* 66, 437–474.
Grindley, N. D. and Reed, R. R. (1985). Transpositional recombination in prokaryotes. *Annu. Rev. Biochem.* 54, 863–896.
Haren, L., Ton-Hoang, B., and Chandler, M. (1999). Integrating DNA: transposases and retroviral integrases. *Annu. Rev. Microbiol.* 53, 245–281.
Scott, J. R. and Churchward, G. G. (1995). Conjugative transposition. *Annu. Rev. Immunol.* 49, 367–397.

21.6 Общие посредники и полупродукты транспозиции

Обзор

Mizuuchi, K. (1992). Transpositional recombination: mechanistic insights from studies of Mu and other elements. *Annu. Rev. Biochem.* 61, 1011–1051.
Pato, M. L. (1989). Bacteriophage mu. In Berg, D. E. and Howe, M., eds., *Mobile DNA.* Washington, DC: American Society for Microbiology Press, pp. 23–52.

Исследование

Aldaz, H., Schuster, E., and Baker, T. A. (1996). The interwoven architecture of the Mu transposase couples DNA synthesis to catalysis. *Cell* 85, 257–269.
Savilahti, H. and Mizuuchi, K. (1996). Mu transpositional recombination: donor DNA cleavage and strand transfer in *trans* by the Mu transpose. *Cell* 85, 271–280.

21.8 Нерепликативная транспозиция сопровождается разрывом и воссоединением

Исследование

Bender, J. and Kleckner, N. (1986). Genetic evidence that Tn10 transposes by a nonreplicative mechanism. *Cell* 45, 801–815.
Bolland, S. and Kleckner, N. (1996). The three chemical steps of Tn10/IS10 transposition involve repeated utilization of a single active site. *Cell* 84, 223–233.
Davies, D. R., Goryshin, I. Y., Reznikoff, W. S., and Rayment, I. (2000). Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. *Science* 289, 77–85.
Haniford, D. B., Benjamin, H. W., and Kleckner, N. (1991). Kinetic and structural analysis of a cleaved donor intermediate and a strand transfer intermediate in Tn10 transposition. *Cell* 64, 171–179.
Kennedy, A. K., Guhathakurta, A., Kleckner, N., and Haniford, D. B. (1998). Tn10 transposition via a DNA hairpin intermediate. *Cell* 95, 125–134.

21.9 В транспозиции TnA участвуют транспозаза и резольваза

Обзор

Sherratt, D. (1989). Tn3 and related transposable elements: site-specific recombination and transposition. In Berg, D. E. and Howe, M. M., eds., *Mobile DNA*. Washington, DC: American Society for Microbiology Press, pp. 163–184.

Исследование

Droge, P. et al. (1990). The two functional domains of gamma delta resolvase act on the same recombination site: implications for the mechanism of strand exchange. *Proc. Natl. Acad. Sci. USA* 87, 5336–5340.

Grindley, N. D. et al. (1982). Transposon-mediated site-specific recombination: identification of three binding sites for resolvase at the *res* sites of γ δ and Tn3. *Cell* 30, 19–27.

Li, W., Kamtekar, S., Xiong, Y., Sarkis, G. J., Grindley, N. D., and Steitz, T. A. (2005). Structure of a synaptic gammadelta resolvase tetramer covalently linked to two cleaved DNAs. *Science* 309, 1210–1215.

21.10 Контроль транспозиция Tn10 осуществляется различными путями

Обзор

Kleckner, N. (1989). Tn10 transposon. In Berg, D. E. and Howe, M. M., eds., *Mobile DNA*. Washington, DC: American Society for Microbiology Press, pp. 227–268.

Kleckner, N. (1990). Regulation of transposition in bacteria. *Annu. Rev. Cell Biol.* 6, 297–327.

Исследование

Roberts, D. et al. (1985). IS10 transposition is regulated by DNA adenine methylation. *Cell* 43, 117–130.

21.12 Контролирующие элементы образуют семейства

Обзор

Fedoroff, N. (1989). Maize transposable elements. In Berg, D. E. and Howe, M. M., eds., *Mobile DNA*. Washington, DC: American Society for Microbiology Press, pp. 375–411.

Fedoroff, N. (2000). Transposons and genome evolution in plants. *Proc. Natl. Acad. Sci. USA* 97, 7002–7007.

Gierl, A., Saedler, H., and Peterson, P. A. (1989). Maize transposable elements. *Annu. Rev. Genet.* 23, 71–85.

Исследование

Benito, M. I. and Walbot, V. (1997). Characterization of the maize Mutator transposable element MURA transposase as a DNA-binding protein. *Mol. Cell Biol.* 17, 5165–5175.

Chandler, V. L. and Walbot, V. (1986). DNA modification of a maize transposable element correlates with loss of activity. *Proc. Natl. Acad. Sci. USA* 83, 1767–1771.

Ros, F. and Kunze, R. (2001). Regulation of activator/dissociation transposition by replication and DNA methylation. *Genetics* 157, 1723–1733.

Singer, T., Yordan, C., and Martienssen, R. A. (2001). Robertson's Mutator transposons in A. thaliana are regulated by the chromatin-remodeling gene decrease in DNA Methylation (DDM1). *Genes Dev.* 15, 591–602.

21.14 Роль мобильных элементов в гибридном дисгенезе

Обзор

Engels, W. R. (1983). The P family of transposable elements in *Drosophila*. *Annu. Rev. Genet.* 17, 315–344.

Engels, W. R. (1989). P elements in *Drosophila*. In Berg, D. E. and Howe, M. M., eds., *Mobile DNA*. Washington, DC: American Society for Microbiology Press, pp. 437–484.

21.15 Р-элементы могут быть активированы в клетках зародышевого пути

Исследование

Laski, F. A., Rio, D. C., and Rubin, G. M. (1986). Tissue specificity of *Drosophila* P element transposition is regulated at the level of mPHK splicing. *Cell* 44, 7–19.

Ретровирусы и ретропозоны

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

- Транспозоны и ретропозоны составляют почти половину генома человека.

- Большая часть повторяющейся ДНК в геномах млекопитающих представлена элементами, принадлежащими к одному семейству, организованными подобно транспозонам и ведущими свое происхождение от продуктов активности РНК-полимеразы III.

- Процессированный псевдоген представляет собой производную, полученную в ходе обратной транскрипции последовательности мРНК.

- У элементов LINES отсутствуют ДКП. Поэтому эти ретропозоны кодируют эндонуклеазу, которая производит надрезы для затравки обратной транскрипции.

22.1 Введение

Транспозиция, в которой в качестве обязательного посредника участвует РНК, уникальна для эукариот. Эта разновидность транспозиции обеспечивается способностью **ретровирусов** вставлять ДНК-копии — провирусы — своего РНК-генома в хромосомы клетки-хозяина. Некоторые транспозоны эукариот по своей общей организации, а также по манере совершать транспозицию с промежуточным продуктом в форме РНК, родственны ретро-провирусам. Как класс, такие элементы называют **ретропозонами** или **ретротранспозонами**. Наипростейшие из таких элементов не обладают собственной транспозиционной активностью, но содержат последовательности, опознаваемые активными элементами как субстрат для транспозиции. Таким образом, элементы, в отношении которых применяется РНК-зависимая транспозиция, ранжируют от собственно ретровирусов (способных свободно инфицировать клетку-хозяина) к последовательностям, кодирующим все функции для перемещения себя по геному с принятием промежуточной формы РНК, вплоть до последовательностей, не способных подкрепить собственное перемещение кодированием соответствующих функций. С транспозонами они имеют ту общую диагностическую черту, что в сайте своего внедрения в геном порождают короткие прямые повторы ДНК-мишени.

Даже в тех геномах, где не были обнаружены активные транспозоны, можно найти следы давних событий транспозиции в виде прямых повторов, фланкирующих рассеянные по геному повторяющиеся последовательности. Некоторые из этих геномных последовательностей наделены особенностями, указывающими на то, что ее предшественником являлась последовательность РНК. По всей видимости, РНК была скопирована в форму двухцепочечной ДНК и в таком виде вставлена в хромосому, подобно транспозону.

Как и любое другое самовоспроизведение, репродуктивный цикл ретровируса или ретропозона непрерывен. Момент «начала» цикла задается произвольным образом, исходя из соображений наблюдателя. Поэтому в нашем видении этих элементов генома преобладают те их формы, в которых они обычно оказываются доступны нашему наблюдению (РИС. 22.1). Впервые ретровирусы были обнаружены в форме инфекционных вирусных частиц, способных передаваться между клетками, и весь внутриклеточный цикл (с фигурирующей в нем двухцепочечной ДНК) можно рассматривать, как мероприятие по воспроизводству РНК-содержащего вируса. Ретропозоны были открыты как компоненты генома, а РНК-формы в основном характеризовались с точки зрения их мРНК-функции. Последовательности ретропозонов мы расцениваем как двухцепочечную геномную ДНК, обладающую мобильностью внутри генома; между клетками эти последовательности не мигрируют.

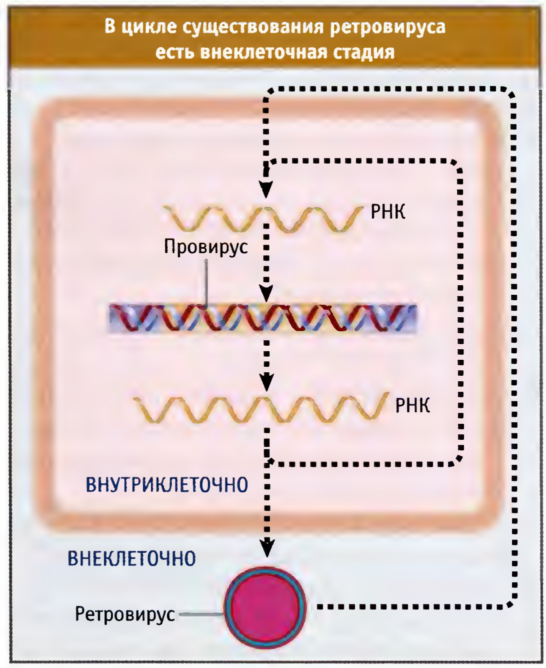


РИС. 22.1 Репродуктивные циклы ретровирусов и ретропозонов можно рассматривать как чередование стадий прямой и обратной транскрипции (копирование ДНК в виде РНК и обратно). Образование инфекционных частиц, покидающих клетку и заражающих другие клетки, свойственно ретровирусам, но не ретропозонам, которые проводят репродуктивные циклы в пределах одной клетки

22.2 Жизненный цикл ретровируса включает в себя стадии, подобные транспозиции

Ключевые понятия:

- Ретровирус содержит две копии генома в форме одноцепочечной РНК.
- Интегрированный провирус представляет собой двухцепочечную ДНК.
- Провирус порождается обратной транскрипцией ретровирусного генома.

Ретровирусы обладают геномами, состоящими из одноцепочечной РНК. Промежуточным продуктом репликации РНК-генома является двухцепочечная ДНК. Жизненный цикл вируса обязательно включает в себя стадию интеграции двухцепочечной ДНК в геном хозяина. Процесс напоминает транспозицию и порождает короткие прямые повторы ДНК-мишени.

Значение этой реакции выходит за пределы потребностей вируса. Вот некоторые из ее последствий:

- Ретровирусная последовательность, интегрированная в геномную ДНК клетки зародышевого пути, остается в клеточном геноме в качестве эндогенного **провируса**. Подобно лизогенному бактериофагу, провирус ведет себя как часть генетического материала организма.
- Клеточные последовательности время от времени рекомбинируют с ретровирусной последовательностью, после чего начинают вместе с ней перемещаться по геному. Эти фрагменты клеточной ДНК могут оказаться вставленными в геном по новым местоположениям.
- Клеточные последовательности, перемещаемые по геному ретровирусом, могут менять свойства клеток, зараженных этим вирусом.

Стадии жизненного цикла ретровируса иллюстрирует РИС. 22.2. Ключевыми моментами являются: конверсия вирусной РНК в ДНК, интеграция (внедрение) этой ДНК в клеточный геном с образованием провируса, транскрипция провируса (и возвращение генетического материала в форму РНК).

Фермент, ответственный за создание исходной ДНК-копии вирусной РНК, называют **обратной транскриптазой**. Фермент конвертирует РНК в линейный дуплекс ДНК в цитоплазме инфицированной клетки. Эта ДНК может быть переведена и в кольцевые формы, но они, по-видимому, в репродукции вируса не участвуют.

Линейная ДНК проделывает свой путь в ядро. Одна или несколько ее копий становятся интегрированными в геном хозяина. За интеграцию отвечает один-единственный фермент, называемый **интегразой**. Клеточные системы транскрибируют провирус с образованием вирусных РНК, используемых двумя способами: и как мРНК, и как генетический материал для упаковки в вирионы. Интеграция, которая для транскрипции необходима, представляет собой нормальную стадию жизненного цикла.

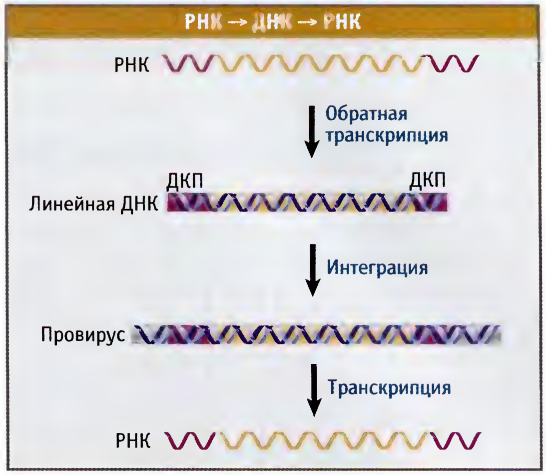


РИС. 22.2 В жизненном цикле ретровируса происходит обратная транскрипция его РНК-генома в двухцепочечную ДНК, которая вставляется в клеточный геном, чтобы, в свою очередь, быть транскрибированной в РНК. ДКП — длинный концевой повтор

В каждый вирион упаковывается по две копии РНК-генома вируса, так что фактически каждая вирусная частица диплоидна. Если клетка оказалась заражена одновременно двумя похожими, но не одинаковыми вирусами, могут образоваться гетерозиготные вирусные частицы, несущие по одной копии генома каждого типа. Диплоидное состояние, возможно, помогает вирусу в захвате фрагментов клеточной ДНК. Ферменты обратная транскриптаза и интеграза переносятся в составе вирусной частицы вместе с РНК-геномом.

22.3 Ретровирусные гены кодируют полипротеины

Ключевые понятия:

- Типичный ретровирус имеет три гена: *gag, pol* и *env*.
- Белки Gag и Pol транслируются с полноразмерного транскрипта вирусного генома.
- Трансляция Pol требует от рибосомы сдвига рамки.
- Белок Env транслируется с отдельной мРНК, образующейся в результате сплайсинга.
- Каждый из трех полипептидных продуктов подвергается процессингу протеазами, с образованием множества белковых продуктов.

Как правило, нуклеотидная последовательность ретровируса содержит три или четыре «гена». В контексте этой главы термин *ген* подразумевает кодирующую область, дающую, благодаря реакциям процессинга, начало множеству белков. Типичный трехгенный ретровирусный геном, организованный в виде последовательности *gag-pol-env*, изображен на РИС. 22.3.

Ретровирусные мРНК имеют обычную структуру: они кэпированы по 5′-концу и полиаденилированы по

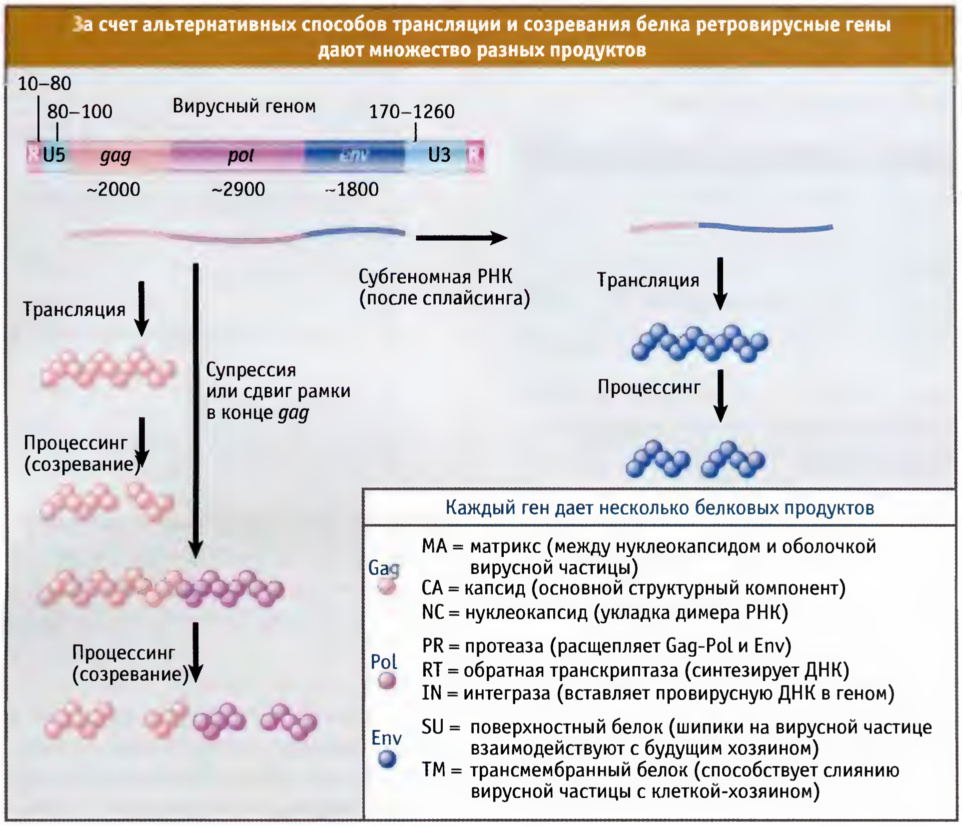


РИС. 22.3 При экспрессии ретровирусных генов получаются полипротеины, которые посттрансляционно доводятся до конкретных функциональных белковых продуктов

3′-концу. Полноразмерная мРНК транслируется с образованием полипротеинов Gag и Pol. Продукт трансляции под названием Gag образуется при прочтении рибосомой отрезка матрицы от инициирующего кодона до первого терминирующего кодона. Чтобы состоялась экспрессия Pol, рибосома должна этот терминирующий кодон проигнорировать.

Разные вирусы пользуются разными способами преодоления терминирующего кодона *gag* как препятствия для трансляции, в зависимости от того, как соотносятся между собой открытые рамки считывания *gag* и *pol*. Если вторая рамка представляет собой продолжение первой, то механизм подавления терминации, приводящий к образованию единого полипептида Gag-Pol, основывается на узнавании терминирующего кодона молекулой глютамил-тРНК. Если же *gag* и *pol* не попадают в одну рамку, соответствующий сдвиг транслируемой матрицы (рибосомальный сдвиг рамки) — с тем же эффектом образования единого полипептида Gag-Pol — осуществляет сама рибосома. Доля транслированных прочтением через терминирующий кодон полноразмерных транскриптов обычно составляет ~5%, так что продуцируемое количество белка Gag превышает количество белка Gag-Pol примерно в 20 раз.

Полипротеин Env синтезируется другим путем: сплайсинг полноразмерного транскрипта порождает укороченную *субгеномную* мРНК, продуктом трансляции которой и является Env.

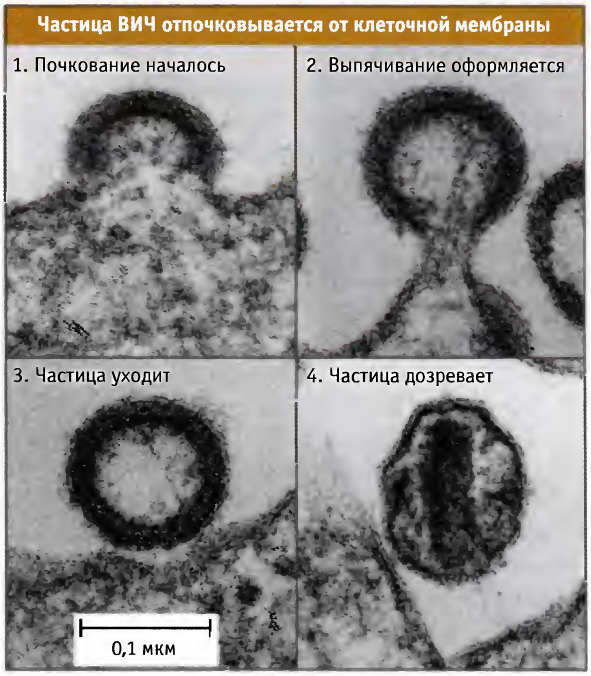


РИС. 22.4 Ретровирусы ВИЧ (от *англ.* human immunodeficiency virus — вирус иммунодефицита человека) отпочковываются от цитоплазматической (наружной) мембраны инфицированной клетки. Фотография любезно предоставлена Matthew A. Gonda, Ph.D., Chief Executive Officer, IntePHKtional Medical Innovations, Inc.

Ген *gag* несет в себе информацию о белковых компонентах нуклеопротеиновой сердцевины вириона. В гене *pol* закодировано энзиматическое обеспечение функций синтеза и рекомбинации нуклеиновых кислот. Ген *env* кодирует компоненты оболочки вирусной частицы. В образовании оболочки, помимо сугубо вирусных *env*-белков, участвуют секвестрированные компоненты клеточной цитоплазматической мембраны.

Оба Gag или Gag-Pol и Env — это полипротеины (множественные белки). Протеаза расщепляет их на обособленные белки, обнаруживаемые в зрелых вирионах. Варианты кодирования белка-носителя протеазной активности у разных вирусов различны: он может представлять собой часть Gag, или часть Pol, или вообще продукт дополнительной независимой открытой рамки считывания.

Продукция инфекционной ретровирусной частицы включает в себя упаковку РНК в сердцевину, окружение ее капсидными белками, отделение в составе сегмента цитоплазматической мембраны клетки-хозяина. Этапы процесса показаны РИС. 22.4. Инфекция представляет собой процесс, обратный высвобождению инфекционных частиц: новая клетка заражается вирусом посредством слияния вириона с цитоплазматической мембраной и высвобождения его содержимого во внутриклеточную среду.

22.4 ДНК ретровируса является продуктом обратной транскрипции

Ключевые понятия:

- РНК ретровируса несет концевые повторы короткой последовательности R; нуклеотидные последовательности 5'- и 3'-концевых областей вирусной РНК называют R-U5 и U3-R, соответственно.
- Обратная транскриптаза начинает синтезировать ДНК, если затравка в форме тРНК связывается с сайтом РНК ретровируса, отстоящим от 5'-конца на расстояние от 100 до 200 оснований.
- Когда фермент доходит до конца матрицы, 5'-концевые основания РНК становятся предметом деградации, что открывает 3'-конец ДНК-продукта.
- Основания открытого 3'-конца комплементарно взаимодействуют с 3'-концом другой молекулы РНК ретровируса.
- Синтез ДНК продолжается; его двухцепочечный конечный продукт имеет повторяющие друг друга концевые области, а именно U3-R-U5.
- Во второй раз смена матрицы обратной транскриптазой, обусловленная переходом на новую матрицу синтезируемой цепи ДНК, происходит, когда фермент использует собственный ДНК-продукт как матрицу для синтеза второй цепи дуплекса.
- Переход синтезируемой цепи ДНК на другую матрицу — ключевое событие рекомбинации по механизму выбора копии.

Ретровирусы называют **вирусами (+)-цепи**, потому что вирусная РНК (первичный агент инфекции) сама кодирует белковые продукты. Фермент обратная транскриптаза отвечает за перевод вирусного генома в комплементарную ему цепь ДНК, называемую **(–)-цепью ДНК**. Последующие стадии продукции ДНК-дуплекса также катализируются обратной транскриптазой. Обладание собственной ДНК-полимеразной активностью позволяет этому ферменту синтезировать вторую цепь ДНК на матрице одноцепочечного обратного транскрипта. Вторую цепь ДНК-дуплекса называют **(+)-цепью ДНК**. Все та же обратная транскриптаза обладает еще и активностью РНКазы Н, которая, в качестве необходимого дополнения к ДНК-полимеразной активности, вызывает деградацию РНК в составе гибрида с ДНК. Все обратные транскриптазы ретровирусов, называемые также ревертазами, проявляют значительное сходство своих аминокислотных последовательностей; соответствующие черты гомологии можно выявить и у некоторых ретропозонов.

На РИС. 22.5 приведено сравнение первичных структур ДНК и РНК ретровируса. РНК ретровируса с обеих сторон оканчивается прямыми повторами, называемыми **R-сегментами**. Их длина колеблется в пределах от 10 до 80 нуклеотидов. 5'-концевую последовательность вирусной РНК называют R-**U5**, 3'-концевую последовательность называют **U3**-R. R-сегменты используются для образования более протяженных прямых концевых повторов при переводе нуклеотидной последовательности ретровируса из РНК-формы в ДНК-форму. Длинные концевые повторы (ДКП) обнаруживаются у линейной двухцепочечной вирусной ДНК (РИС. 22.6 и РИС. 22.7). Укороченность интегрированной формы вируса на 2 пн с каждого конца — следствие специального механизма интеграции (см. рис. 22.9).

Как и другим ДНК-полимеразам, обратной транскриптазе (ревертазе) требуется затравка. Естественной затравкой служит тРНК. Каждый вирион содержит «незаряженную» (аминокислотой) тРНК бывшего хозяина. Последовательность из 18 оснований в 3'-концевой области тРНК комплементарно взаимодействует с сайтом вирусной РНК, заглубленным на 100–200 оснований от ее 5'-конца. Одновременно та же молекула тРНК может взаимодействовать с сайтом вблизи 5'-конца другой вирусной РНК, ассистируя таким образом димеризации вирусных РНК-геномов.

Итак, ревертаза начинает синтезировать ДНК и движется в направлении 5'-конца матрицы, от которого она изначально отстоит всего на 100-200 нуклеотидов. Каким же образом ей удастся преобразовать интактный геном ретровируса в форму ДНК? (Этот случай представляет собой экстремальный вариант общей проблемы репликации краев линейной молекулы нуклеиновой кислоты; см. разд. 16.2 «Края линейной ДНК сложно реплицировать»).

В условиях *in vitro* синтез доходит до конца матрицы, порождая короткую молекулу ДНК, называемую strong-stop (–)-цепи ДНК. В клетке *in vivo* таких молекул не обнаруживается, поскольку в естественных условиях синтез непременно продолжается реакцией, проиллюстрированной на рис. 22.6. Ревертаза, переставляя синтезируемую цепь ДНК на новую матрицу, совершает смену матрицы. Это первый из двух прыжков между матрицами, которые фермент совершает за время работы над каждой копией генома ретровируса.

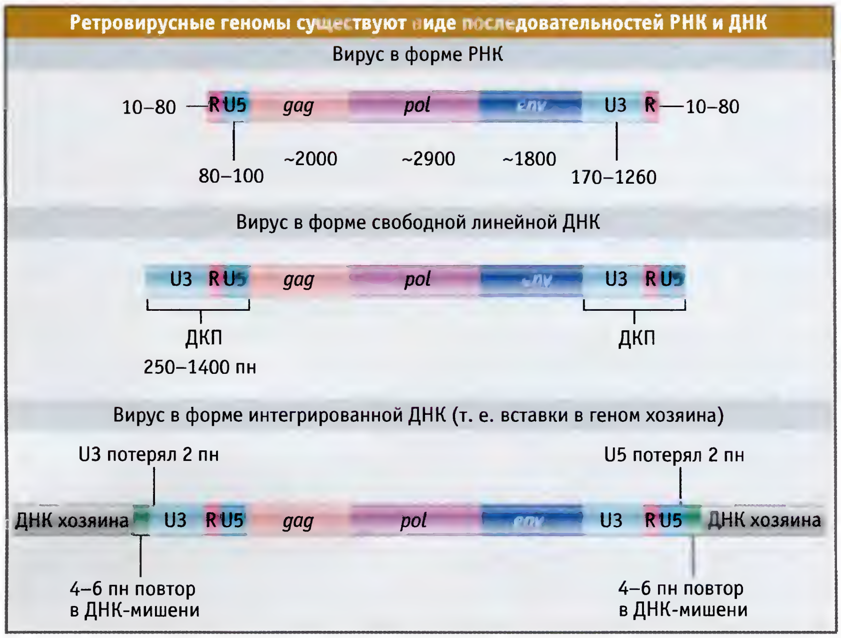


РИС. 22.5 **Ретровирусная РНК оканчивается прямыми повторами (R), свободная линейная вирусная ДНК оканчивается ДКП, провирус оканчивается теми же ДКП, укороченными на 2 пн**

Чтобы смена матрицы могла состояться, ревертаза разрушает 5′-концевой R-сегмент матрицы, используя присущую ей активность РНКазы Н. Открывающийся при этом 3′-концевой одноцепочечный выступ новосинтезированной ДНК, по принципу комплементарности, склеивается с 3′-концевым R-сегментом другой молекулы РНК. После состоявшейся таким образом смены матрицы обратная транскрипция продолжает свой ход через U3-сегмент дальше, во внутренний массив РНК.

3′-концевой R-сегмент РНК, подставляемый под 3′-концевой одноцепочечный выступ ДНК, в принципе, может являться противоположным концом той же самой молекулы РНК (тогда имеет место внутримолекулярное спаривание), но обычно принадлежит другой молекуле (межмолекулярное спаривание). Межмолекулярное спаривание изображено на рис. 22.6; внутримолекулярное спаривание не оставляло бы возможности изменять сайт связывания праймера от цикла к циклу, в то время как эксперименты показывают, что последовательность тРНК-затравки в жизненном цикле ретропозона не наследуется.

Результатом смены матрицы и продолжения синтеза является добавление (–) U5-сегмента к 5′-концевой области (–)-цепи ДНК. Отрезок нуклеотидной последовательности U3-R-U5 (в прочтении по (+)-цепи) — это и есть длинный концевой повтор (ДКП), поскольку позже (+) U3-сегмент будет добавлен к 5′-концу (+)-цепи ДНК, и он обретет вид U3-R-U5. Протяженность ДКП варьирует в пределах от 250 до 1400 пн (см. рис. 22.5).

Процесс синтеза (+)-цепи ДНК с ДКП на противоположном (левом) краю показан на рис. 22.7. Ревертаза затравляет синтез (+)-цепи от остаточного фрагмента деградировавшей РНК-матрицы в гибриде с новосинтезированной (–)-цепью. Strong-stop(+)-цепь ДНК оказывается синтезированной, когда фермент достигает конца матрицы (то есть считывает (–)-цепь ДНК до ее 5′-конца). Потом эта короткая адаптерная (+)-цепь ДНК совершает прыжок на противоположный конец (–)-цепи. Вероятно, ее оттесняет туда продукт второго раунда синтеза (+)-цепи, с другим остаточным фрагментом РНК (на рисунке расположенным левее) в качестве затравки.

Каждая ретровирусная частица (вирион) несет в себе два экземпляра РНК-генома вируса. Это обстоятельство благоприятствует рекомбинации как событию, естественному для жизненного цикла вируса. Рекомбинация может случиться во время синтеза (–)-цепи ДНК и/или во время синтеза (+)-цепи ДНК.

Межмолекулярное спаривание, изображенное на рис. 22.6, позволяет произойти рекомбинации между последовательностями двух РНК, сменяющих друг друга в качестве матриц синтеза (–)-цепи ДНК. У ретровирусов рекомбинация в основном и происходит на этой стадии, когда насцентная цепь ДНК оказывается, в процессе обратной транскрипции, перенесенной с одной матрицы на другую.

(+)-цепь ДНК может быть синтезирована прерывисто, то есть в результате реакции, включающей в себя несколько инициаций синтеза в разных внутренних сайтах. Перенос цепи синтезируемого фрагмента на другую матрицу на этой стадии также возможен, но в действительности случается не так часто.

Сходство двух описанных вариантов в том, что рекомбинация в том и другом случае является результатом смены матрицы в продолжение акта синтеза ДНК. Такой механизм рекомбинации называют **выбором копии**. Многие годы выбор копии рассматривался как возможный механизм общей (гомологичной) рекомбинации.

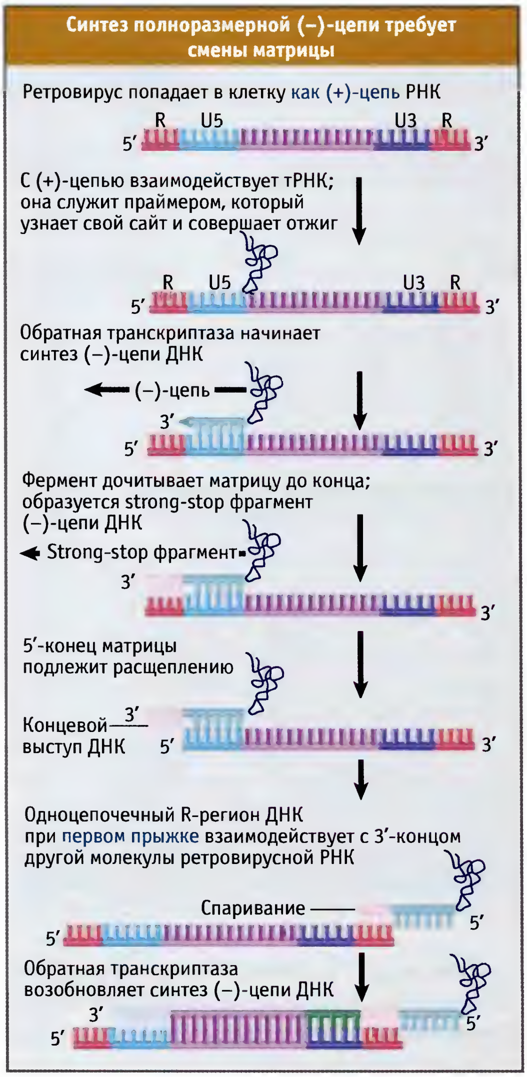


РИС. 22.6 Синтез (–)-цепи идет со сменой матрицы в процессе обратной транскрипции

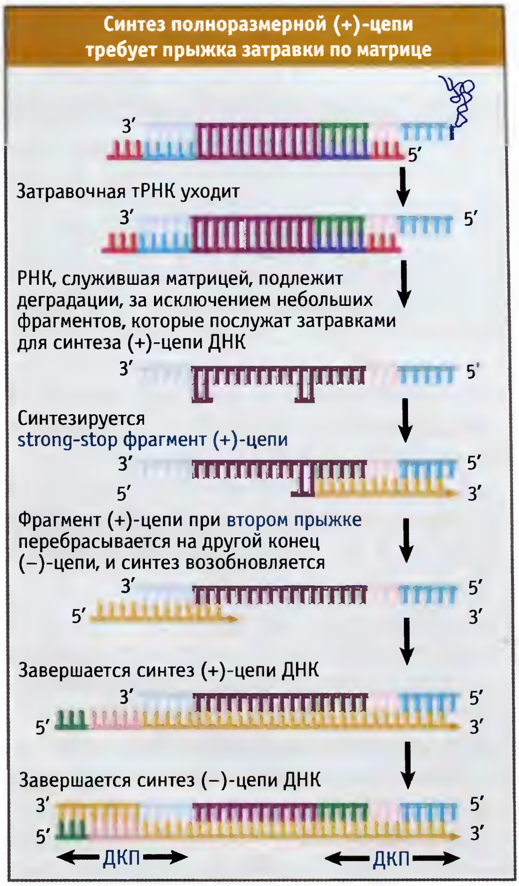


РИС. 22.7 При синтезе (+)-цепи ДНК имеет место второй прыжо затравочного конца

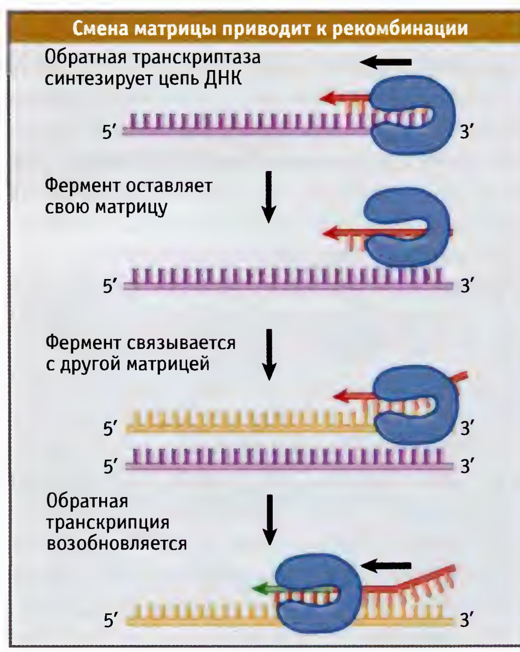


РИС. 22.8 Рекомбинация по типу выбора копии происходит из-з того, что обратная транскриптаза в процессе реакции меняет мат рицу. Для наглядности переход фермента с матрицы на матриц изображен в виде отдельных этапов

Маловероятно, что им на самом деле пользуются клеточные системы, но во время инфекции РНК-содержащими вирусами (в том числе воспроизводящимися сугубо через формы РНК, как, например, вирус полиомиелита) этот механизм рекомбинации весьма актуален.

Переход цепи — дополнительно к вынужденной реакции переноса цепи при «упоре» синтеза в конец матрицы — с определенной частотой происходит в каждом цикле обратной транскрипции. Механизм перехода цепи в данном случае изучен недостаточно, принцип же иллюстрирует РИС. 22.8. В условиях *in vivo* обратная транскрипция происходит внутри нуклеопротеинового комплекса, в составе которого цепь РНК-матрицы связана с компонентами вириона, среди которых — главный белок капсида. В случае ВИЧ добавление этого белка, NCp7, вызывает рекомбинацию в системе *in vitro*. Возможно, эффект непрямой: NCp7 каким-то образом видоизменяет структуру РНК-матрицы, что, в свою очередь, влияет на склонность ревертазы сменять одну матричную цепь на другую.

22.5 Вирусная ДНК интегрируется в хромосому

Ключевые понятия:

- Организация провирусной ДНК в хромосоме такая же, как у транспозона: провирус фланкирован короткими прямыми повторами последовательности сайта-мишени.
- Линейная ДНК вставляется непосредственно в хромосому хозяина при участии ретровирусного фермента интегразы.
- В процессе интеграции последовательность ДНК ретровируса теряет по две пары оснований с каждого конца.

Структурная организация у внедрившегося провируса почти такая же, как и у линейной ДНК. Длинные концевые повторы (ДКП) провируса идентичны между собой. 3′-концевая область U5-сегмента ДКП представляет собой инвертированную 5′-концевую область U3-сегмента, так что каждый ДКП (равно как и провирус в целом) окаймлен короткими инвертированными концевыми повторами. Внедренный провирус подобен транспозону: оканчивается инвертированными повторами и фланкирован короткими прямыми повторами последовательности-мишени.

Провирус образуется как непосредственная вставка линейной ДНК вируса в сайт-мишень. Кроме линейной, существуют также кольцевые формы ДНК вируса. Одна из кольцевых форм несет два ДКП, расположенных вплотную друг к другу в результате склеивания концов линейной ДНК. Большинство же кольцевых ДНК вируса несет всего одну копию последовательности ДКП — вероятно, продукт события рекомбинации (а может быть, внутримолекулярного спаривания?). Кольцевые молекулы ДНК ретровирусов долгое время рассматривались как возможные промежуточные продукты интеграции (по аналогии с интеграцией ДНК бактериофага лямбда) до тех пор, пока не стало известно, что в интеграции ретровирусов используется линейная форма ДНК.

Интеграцию линейной ДНК катализирует интеграза — белок вирусного происхождения. Интеграза влияет и на линейную ДНК ретровируса, и на клеточную ДНК-мишень. Реакцию иллюстрирует рис. 22.9.

Для успеха этой реакции особенно важна первичная структура концов ДНК ретровируса — подобно тому, как мутации в концевых областях транспозона препятствуют его интеграции. Самая консервативная особенность — присутствие динуклеотида СА близко к наружному концу каждого из коротких инвертированных повторов, окаймляющих ДНК ретровируса (и каждый ее ДКП в отдельности). Интеграза, в составе нуклеопротеинового комплекса, сводит вместе тупые концы линейной ДНК, а затем, удаляя нуклеотиды по внешнему краю консервативных СА, превращает тупые двухцепочечные концы в усеченные, с 5′-выступом и 3′-заглублением. В общем случае, ДНК при этом теряет по два нуклеотидных основания с каждого конца.

Выбор сайтов мишеней, в плане их нуклеотидной последовательности, достаточно случаен. Интеграза вносит в сайт-мишень надрезы со смещением. В примере на

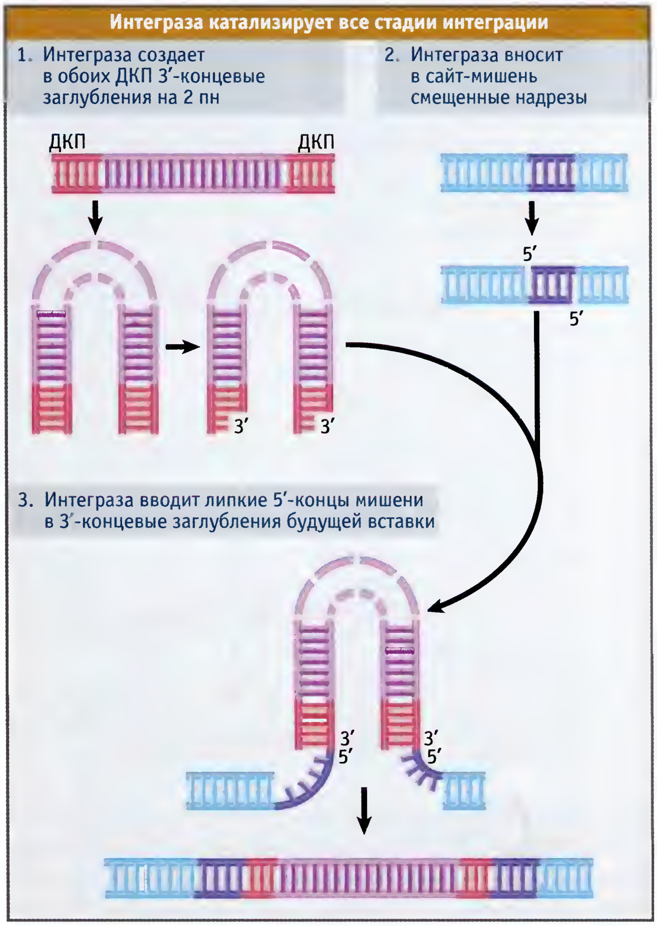


РИС. 22.9 Интеграза — это единственный вирусный белок, необходимый для интеграции, при которой оба ДКП теряют по две крайних нуклеотидных пары и вставляются между 4 пн повторами последовательности сайта-мишени

рис. 22.9 смещение (область которого в ходе реакции преобразуется в прямые повторы сайта-мишени) составляет 4 пн. Протяженность повтора мишени специфична для каждого вируса и может равняться 4, 5 или 6 пн. По всей видимости, она определяется стереохимическими параметрами взаимодействия интегразы с ДНК-мишенью.

5′-концы надрезов ДНК-мишени ковалентно присоединяются к заглубленным 3′-концам ДНК ретровируса. В этот момент линейная двухцепочечная ДНК ретровируса оказывается соединенной с сайтом-мишенью одной цепью с каждого конца. Заполнением одноцепочечных участков мишени занимаются клеточные ферменты репарации, и в ходе этой реакции 5′-концевые выступы ДНК ретровируса подлежат удалению. Таким образом, внедрение в клеточный геном укорачивает ДНК ретровируса на 2 пн с каждой стороны; это соответствует потере 2 пн 5′-концевого U3-сегмента и 2 пн 3′-концевого U5-сегмента.

Итак, ДНК ретровируса интегрирует в геном хозяина в произвольном месте. Успешно инфицированная клетка получает в свой геном от одной до десяти копий провируса. (Заражение клетки инфекционными частицами

происходит, конечно же, через цитоплазму, а интеграция линейной ДНК вируса в геном хозяина происходит в ядре. Ретровирусы могут размножаться только в пролиферирующих (активно делящихся) клетках, потому что попадание в ядро чужеродной ДНК возможно только при митозе, и именно на этой стадии клеточного цикла генетический материал вируса (в форме линейной ДНК) получает доступ к содержимому клеточного ядра.

В U3-сегменте каждого ДКП содержится промотор транскрипции. Промотор в левом ДКП (для (+)-цепи он 5′-концевой) отвечает за инициацию транскрипции провируса. Если учесть, что несущий его U3-сегмент появляется в левом ДКП благодаря смене матрицы ферментом-ревертазой в процессе синтеза будущей ДНК провируса, то получается, что этот промотор как таковой был порожден самим переводом РНК вируса в форму двухцепочечной ДНК.

Иногда (довольно редко) промотор в составе правого ДКП оказывает поддержку транскрипции геномных последовательностей хозяина, смежных с сайтом интеграции. ДКП содержит также последовательность-энхансер, активирующую близлежащие промоторы — как вирусные, так и клеточные. Через активацию определенных генов интеграция ретровируса может стать причиной перехода клетки в опухолеродное состояние. Возможна ли эксцизия из генома внедренных в него провирусов? Да, благодаря гомологичной рекомбинации, которая может произойти между ДКП провируса; одиночные ДКП, оставшиеся после эксцизии, можно обнаружить во многих клеточных геномах.

До сих пор мы рассматривали ретровирусы с точки зрения их инфекционного цикла, где этап интеграции необходим для последующей наработки копий РНК. Следует заметить, что, когда ДНК ретровируса интегрирует в геном клетки зародышевого пути, она становится наследуемым эндогенным провирусом организма. Обычно эндогенные провирусы не экспрессируются, но их могут активировать неординарные события внешнего рода, такие как инфекция клетки другим вирусом.

22.6 Ретровирусы могут переносить клеточные последовательности

Ключевые понятия:

- Трансформирующие ретровирусы возникают в результате рекомбинации, при которой последовательность клеточной РНК заменяет часть последовательности РНК ретровируса.

В свете того обстоятельства, что ретровирусы как класс способны к захвату и преобразованию клеточных последовательностей, интерес к их жизненному циклу весьма велик. **Трансдуцирующий вирус** способен перенимать клеточную последовательность, с частичной заменой собственной последовательности геном *v-onc*, изображен на РИС. 22.10. Вместо обычных белков Gag, Pol и Env у него синтезируется белок Gag-v-Onc. Такой вирус **репликационно-дефектен**: не самодостаточен в плане прохождения инфекционного цикла, нуждаясь в **вирусе-помощнике**, который компенсирует ему отсутствие собственных энзиматических функций.

Название *onc* представляет собой сокращение слова **«онкогенез»** — способность *трансформировать* культуры клеток, упраздняя ограничения на их рост, что предоставляет клеткам неограниченную способность к делению. Последовательность *onc* может принадлежать любому из многочисленных генов, чья экспрессия в форме *v-onc* приводит к трансформации клетки в опухолеродный фенотип. Причиной опухолевой трансформации может стать как клеточный, так и вирусный ген *onc*.

Наличие у вируса гена *v-onc* придает ему способность к трансформации клеток-хозяев определенного типа. У хозяина есть локусы с гомологичными последовательностями — генами *c-onc*. Каков механизм присвоения генов *onc* ретровирусами? Ключом к ответу служит расхождение в первичных структурах *c-onc* и *v-onc*. Кодирующая последовательность гена *c-onc* обычно прерывается интронами, которых нет у *v-onc*. Это указывает на происхождение генов *v-onc* от транскриптов *c-onc*, прошедших сплайсинг.

Модель возникновения трансформирующих вирусов проиллюстрирована РИС. 22.11. Допустим, возле гена *c-onc* в клеточный геном внедрился ретровирус. За внедрением последовала делеция, в результате которой провирус и *c-onc* слились. Транскрипция области их слияния порождает РНК-химеру, часть которой соответствует последовательности вируса, а другая часть — последовательности *onc* хозяина. Интроны, входящие в состав обеих частей, удаляются при сплайсинге. Получившаяся РНК обладает сигнальной последовательностью, необходимой для упаковки ее в вирион, и она будет туда упакована, если в клетке есть другая, интактная, копия провируса. Тогда некоторые вирионы (диплоидные частицы)



РИС. 22.10 У трансформирующего вируса с неполноценным обеспечением репликации часть собственной нуклеотидной последовательности заменена фрагментом последовательности клеточного генома. Такой дефектный вирус может реплицироваться при участии вируса-помощника — носителя всех недостающих функций

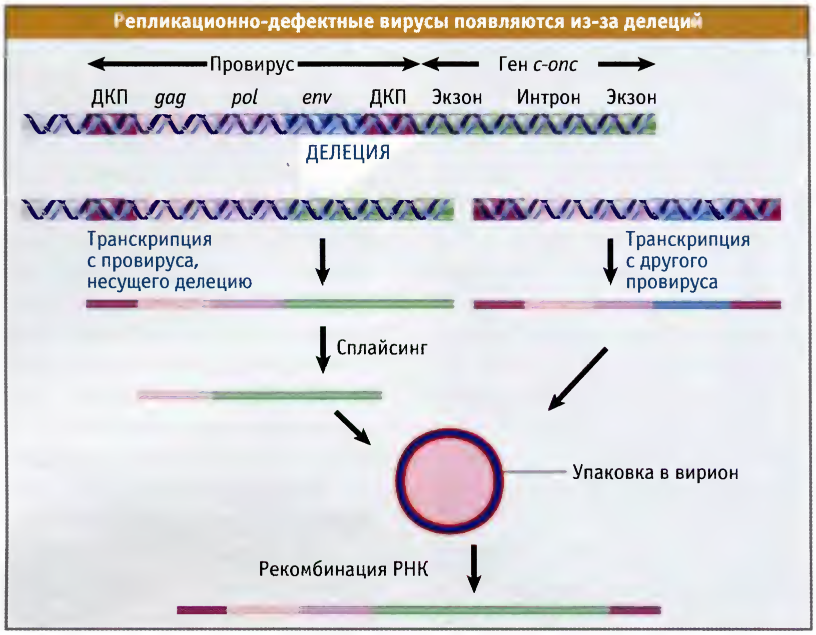


РИС. 22.11 Вирусы с неполноценным обеспечением репликации могут возникать из-за крупных делеций, затрагивающих провирус и приводящих к слиянию его нуклеотидной последовательности с последовательностью клеточного генома. Аномальный транскрипт такого провируса может быть упакован в одну частицу с нормальным вирусным РНК-геномом. Если между двумя РНК произойдет негомологичная рекомбинация, в следующем инфекционном цикле возможно появление трансформирующего вируса с неполноценным обеспечением репликации

из тех, что продуцирует эта клетка, будут содержать по одной копии РНК-химеры и РНК вируса.

Продукт рекомбинации между последовательностями двух этих РНК, с обеих сторон оканчивающийся повторами вирусного происхождения, может обладать трансформирующим потенциалом. По ходу инфекционного цикла ретровируса рекомбинация происходит разнообразными способами и с высокой частотой. Предположительно, негомологичная реакция между РНК вируса и клеточной частью РНК-химеры подчиняется одному из механизмов рекомбинации, используемых вирусами.

Общность черт ретровирусов (уникальных для них как класса) говорит о возможности происхождения ретровирусов от общего предка. Как если бы какие-то первичные вставочные последовательности (IS-элементы) окружили клеточный ген, кодирующий полимеразу; состоявшийся таким образом модуль имел бы вид ДКП-*pol*-ДКП. Эволюции в направлении превращения в инфекционный вирус должно было сопутствовать усовершенствование кодируемых вирусом манипуляций с ДНК- и РНК-субстратами, в том числе обретение генов, чьи продукты смогли бы упаковывать РНК. Другие свойства, например способность к привлечению в свой состав генов с трансформирующей функцией, могли быть приобретены позднее. Нет оснований полагать, что механизм присвоения вирусом клеточных последовательностей уникален для генов *onc*; вирусы, несущие гены *onc*, могут обладать селективным преимуществом благодаря стимулирующему эффекту в отношении клеточного роста.

22.7 *Ty*-элементы дрожжей аналогичны ретровирусам

Ключевые понятия:

- По своей организации *Ty*-элементы похожи на эндогенные ретровирусы.
- *Ty*-элементы — это ретропозоны с собственной ревертазной активностью, промежуточный продукт их транспозиции представлен РНК.

Ty-элементы — это представители семейства повторяющихся последовательностей ДНК, рассеянных в разных местоположениях в геномах разных штаммов дрожжей. (Обозначение *Ty* представляет собой сокращение от transposon yeast, дрожжевой транспозон.) Характерный след события транспозиции — повтор последовательности мишени с каждой стороны от внедрившегося элемента — в данном случае составляет 5 пн. *Ty*-элементы — это **ретропозоны**, по механизму транспозиции подобные ретровирусам. Частота транспозиций *Ty* ниже, чем у большинства бактериальных транспозонов, и составляет ~10^{-8}–10^{-7}.

Отдельные *Ty*-элементы могут различаться между собой довольно сильно. Большинство принадлежит к двум основным классам, называемым *Ty1* и *Ty917*. Общий принцип организации одинаков для этих классов, он изображен на РИС. 22.12. Каждый элемент имеет длину 6,3 тпн и с каждой стороны оканчивается 330 пн прямого повтора, называемого δ. Отдельно взятый *Ty*-элемент может иметь

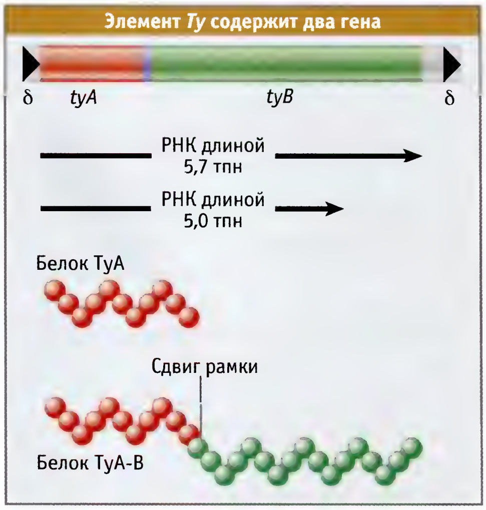


РИС. 22.12 Элементы *Ty* оканчиваются короткими прямыми повторами и при транскрипции дают две перерекрывающиеся РНК. Они содержат две открытые рамки считывания, соответствующие ретровирусным генам *gag* и *pol*

множество отличий от прототипа своего класса, таких как нуклеотидные замены, вставки или делеции. Типичный дрожжевой геном содержит ~30 копий элементов класса *Ty1* и ~6 копий элементов класса *Ty917*. Кроме них в том же геноме присутствуют ~100 независимых *дельта*-элементов, также называемых «сольными δ».

Последовательности δ гетерогенны, хотя в отдельно взятом элементе два концевых повтора идентичны, во всяком случае, очень похожи. Последовательности δ, ассоциированные с *Ty*-элементами, более консервативны, чем сольные δ, что указывает на важность точного соблюдения структуры повторов для их опознавания ферментом транспозиции.

Продукты транскрипции *Ty*-элемента, поли(А)$^+$ РНК двух видов, в сумме составляют >5% тотальной мРНК гаплоидной дрожжевой клетки. Транскрипты обоих видов инициируются внутри промотора, расположенного в левом δ-повторе *Ty*-элемента. Одни из них терминируются спустя 5 тпн; другие — спустя 5,7 тпн внутри правого δ-повтора.

Последовательность *Ty*-элемента содержит две открытых рамки считывания. Эти рамки направлены в одну и ту же сторону, но считываются в разных фазах, несмотря на то что имеют область перекрывания в 13 аминокислот. Судя по последовательности, первая рамка, *TyA*, кодирует ДНК-связывающий белок. Последовательность *TyB* содержит области гомологии с ревертазой, протеазой и интегразой ретровирусов.

По своей организации и функциональному предназначению *TyA* и *TyB* аналогичны *gag* и *pol* ретровирусов. Открытые рамки считывания *TyA* и *TyB* экспрессируются в двух формах. Белок TyA отображает рамку *TyA* и заканчивается вместе с ней. Рамка *TyB* может быть экспрессирована только как часть соединенного полипептида, в котором область, соответствующая рамке *TyA*, слита с областью, соответствующей рамке *TyB*, в результате высокоспецифичного сдвига рамки, позволяющего трансляции преодолеть кодон-терминатор. Аналогичные события сопутствуют трансляции *gag-pol* у ретровирусов.

Рекомбинация между *Ty*-элементами, по всей видимости, проходит в режиме вспышек геномного масштаба: если в клетке зафиксировано одно такое событие, то велика вероятность обнаружения других. Между двумя *Ty*-элементами с разной локализацией может произойти конверсия, и тогда один из них оказывается замещенным последовательностью другого.

Гомологичная рекомбинация между прямыми δ-повторами может привести к эксцизии *Ty*-элемента. Множество сольных δ-элементов может представлять собой следы событий эксцизии. Эксцизия может быть сопряжена с реверсией эффекта мутации от встраивания *Ty*; степень реверсии может определяться тем, какую именно последовательность ретропозон δ оставил после себя.

Парадоксально, что хотя оба δ-повтора имеют одинаковую последовательность, в δ на одном краю *Ty*-элемента активен промотор, а в δ на другом краю активен терминатор. Так бывает и у других мобильных элементов, в том числе у ретровирусов.

Ty-элементы — это классические ретропозоны: формой перемещения по геному им служит РНК. Оригинальный экспериментальный протокол, используемый для исследования особенностей транспозиции *Ty*, проиллюстрирован на РИС. 22.13. *Ty*-элемент, несущий

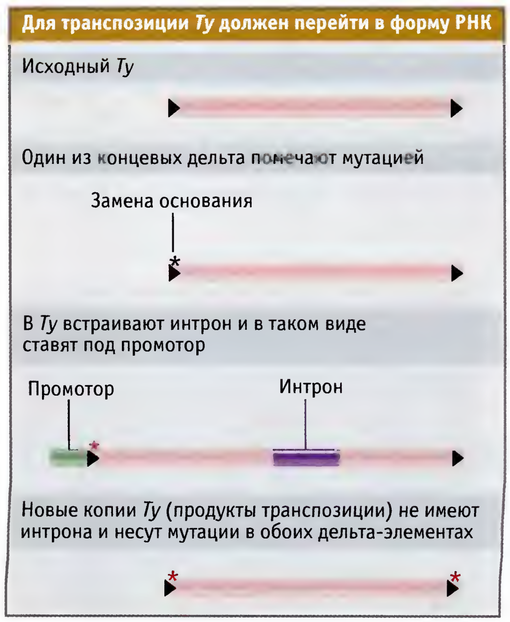


РИС. 22.13 Элемент *Ty* с уникальной искусственной конструкцией использовался для доказательства, что *Ty* являются ретропозонами (то есть что для транспозиции такому элементу необходимо перейти в форму РНК). Новые копии *Ty* оказывались лишенными интрона (то есть в процессе транспозиции подвергались сплайсингу, характерному для РНК, но не для ДНК) и обладали идентичными концевыми повторами, происходящими от одного из концевых повторов оригинального конструкта

искусственную вставку, содержащую интрон, помещают в плазмиду под контроль промотора GAL. Эту плазмиду вносят в дрожжевые клетки. Вскоре в клеточном геноме появляются многочисленные копии *Ty*-элемента, меченного искусственной вставкой — продукты событий транспозиции из плазмиды в хромосому. Эти копии, в отличие от *Ty*-элемента в составе плазмиды, не содержат интрона.

Известен только один способ удаления интронов: сплайсинг РНК. Судя по всему, транспозиция *Ty* следует тому же механизму, каким пользуются ретровирусы. *Ty*-элемент транскрибируется в форму РНК, которую опознает клеточный аппарат сплайсинга. После сплайсинга РНК узнается обратной транскриптазой и «восстанавливается» в форму двухцепочечной ДНК.

Аналогия с ретровирусами заходит дальше. Если у *Ty*-элемента в плазмиде последовательности δ-повторов между собой отличаются, его копии, попадающие в клеточный геном в результате транспозиции, обладают одинаковыми δ-повторами, идентичными 5′-концевому повтору исходного элемента. Если отождествить δ-повтор с ДКП ретровируса, состоящим из сегментов U3-R-U5, то промежуточный продукт транспозиции, *Ty*-РНК, содержит область между двумя R-сегментами, включительно. Полный δ-ДКП восстанавливается при добавлении U5-сегмента к 3′-концу и U3-сегмента к 5′-концу, в точности как показано на рис. 22.3–22.6 для ретровирусов.

Транспозиция *Ty* контролируется генами, расположенными в самом элементе. Промотор GAL, используемый для контроля транскрипции меченного искусственной вставкой *Ty*-элемента в составе плазмиды, является индуцибельным: он включается при добавлении галактозы. Индукция промотора имеет два эффекта. Она необходима, чтобы активировать транспозицию *Ty*-элемента из плазмиды в геномную ДНК дрожжевой клетки; кроме того, активация плазмидного промотора учащает транспозицию



РИС. 22.14 Элементы *Ty* дают начало вирусоподобным частицам. Воспроизведено по *J. Mol. Biol.*, vol. 292, AL-Khayat, H. A., et al., Yeast Ty retrotransposons ..., pp. 65–73. Авторские права 1999, с разрешения Elsevier. Фотография любезно предоставлена Dr. Hind A. AL-Khayat, Imperial College London, United Kingdom

Ty-элементов, уже пребывающих в составе геномной ДНК хозяина. Это говорит о том, что белковые продукты *Ty*-элемента могут *транс*-активно воздействовать на другие подобные элементы (на их РНК).

Хотя *Ty*-элементы не порождают инфекционных частиц, внутри клеток, где была индуцирована их транспозиция, происходит накопление вирусоподобных частиц. Частицы (которые можно разглядеть на РИС. 22.14) содержат полноразмерную РНК, двухцепочечную ДНК, ревертазу и продукт *TyB*, обладающий активностью интегразы. Первичный продукт *TyA*, подобно полипротеину Gag, подлежит расщеплению с образованием зрелых белков сердцевины частицы. Тем самым аналогия между *Ty*-элементом и ретровирусом становится очевидна. *Ty*-элемент ведет себя как ретровирус, который потерял ген *env* и вследствие чего оказался неспособен упаковать свой геном надлежащим образом.

Не все *Ty*-элементы активны: в любом дрожжевом геноме большинство из них, подобно инертным эндогенным провирусам, не занимается распространением собственных копий. Впрочем, у таких элементов сохраняются δ-повторы, благодаря которым они могут стать объектом транспозиции силами белков, обеспечиваемых активным *Ty*-элементом.

22.8 В числе мобильных элементов *Drosophila melanogaster* есть ретропозоны

Ключевые понятия:

- Элемент *copia* — это ретропозон, в изобилии присутствующий у *D. melanogaster*.

О присутствии мобильных элементов в геноме дрозофилы, как и о вставочных последовательностях у бактерий *E. coli*, впервые заподозрили в связи с появлением нестабильных мутаций. У их носителей наблюдалось спонтанное возвращение к дикому типу либо спонтанные делеции одного из флангов сайта мутации. Оказалось, что нестабильные мутации вызываются мобильными последовательностями нескольких типов, с которыми можно ознакомиться по РИС. 22.15. В их числе ретропозоны *copia*, семейство FB, а также P-элементы (обсуждавшиеся в разд. 21.14 «Роль мобильных элементов в гибридном дисгенезе»).

Семейство *copia* изучено лучше всех других ретропозонов. В названии семейства отражено присутствие в геноме большого числа принадлежащих к нему близкородственных последовательностей, кодирующих обильный пул мРНК. Семейство *copia* рассматривается исследователями в качестве парадигмы (это означает, что его особенности, по умолчанию, то есть при отсутствии экспериментальных опровержений, распространяются на несколько других семейств ретропозонов, чьи последовательности не похожи на *copia*, но план строения и тенденции, судя по всему, аналогичны).



РИС. 22.15 Мобильные элементы, встречающиеся в геноме *D. melanogaster*, имеют три типа организации

Численность элементов *copia* в геноме зависит от линии, к которой принадлежит муха; обычно их насчитывается от 20 до 60. Расположение широко рассеянных по геному элементов *copia* подчинено паттернам, которые у разных линий мух частично совпадают.

Различия в паттернах рассеяния мобильных элементов по геному возникали в течение периодов эволюции. При сравнении линий, дивергировавших недавно (за последние лет сорок) в результате своего распространения в лабораторной практике, различий обнаруживается мало. Мы не можем, таким образом, оценить скорость изменений, но молекулярная природа событий, которые обусловливают эти изменения, может быть изучена на клетках, выращиваемых в культуре. Такие клетки содержат в несколько раз больше элементов *copia*, чем до перевода в культуральные условия. Изменения представляют собой вставки последовательностей *copia* в новых сайтах. При адаптации клеток к условиям культивирования частота транспозиций каким-то неизвестным образом временно повышается до уровня в 10^{-4}–10^{-3} событий на поколение.

Элемент *copia* имеет длину ~5000 пн и несет прямые концевые повторы 276 пн. Каждый из этих прямых повторов сам по себе с двух сторон оканчивается инвертированными повторами. В сайте внедрения *copia* образуется прямой повтор 5 пн мишени. Дивергенция между представителями семейства *copia* небольшая, < 5%; варианты последовательностей нередко содержат небольшие делеции. В целом эти особенности присущи и другим семействам ретропозонов, хотя их члены проявляют более высокую степень дивергенции.

Взаимная идентичность двух прямых повторов в составе каждого элемента *copia* означает одно из двух: либо их взаимодействие допускает коррекцию, либо (вероятнее всего) при транспозиции они оба берут свое начало от одного из прямых повторов элемента-предшественника. Как и в случае с *Ty*-элементами, это указывает на родство *copia* с ретровирусами.

Элементы *copia* в составе генома всегда интактны; обособленные копии их концевых повторов не детектируются, хотя их появления можно было бы ожидать как результата рекомбинации, которая делетировала бы заключенный между повторами материал. В самом деле, элементы *copia* время от времени можно встретить в форме свободной кольцевой ДНК; подобно кольцам ДНК ретровирусов, более длинная форма содержит две копии концевого повтора, а более короткая — всего одну. Также были замечены частицы с содержанием ДНК *copia*.

Последовательность *copia* содержит единственную открытую рамку считывания длиной 4227 пн. Между отдельными частями этой рамки *copia* и последовательностями *gag* и *pol* ретровирусов прослеживается гомология. Примечательно отсутствие гомологии с последовательностями *env* ретровирусов, которые требуются для образования оболочки вирусной частицы, так что вряд ли элементы *copia* способны к порождению вирусоподобных частиц.

Продукты транскрипции *copia* в изобилии присутствуют в клетке в виде зрелых поли(A)$^+$ мРНК, представленных как полноразмернми, так и короткими транскриптами. У этих мРНК одинаковая 5′-область, поскольку все они инициируются в середине одного из концевых повторов элемента. По-видимому, белки, кодируемые элементом *copia*, продуцируются с задействованием механизмов сплайсинга РНК и расщепления полипротеинов.

Прямые доказательства транспозиции элементов *copia* в соответствии с тем или иным механизмом отсутствуют; вместе с тем они так сильно напоминают ретровирусов, что неизбежно должны были произойти именно от них. Какие из функций ретровирусов им свойственны, сказать трудно. Нам, конечно же, известно, что они совершают транспозиции, но (как и в случае *Ty*-элементов) ничто не говорит об их способности к инфицированию.

22.9 Систематика ретропозонов

Ключевые понятия:

- К надсемейству вирусных ретропозонов относят транспозоны, мобильность которых реализуется через РНК, не входящую в состав инфекционной частицы.
- Некоторые ретропозоны используют ДКП и тем самым демонстрируют непосредственное сходство с ретровирусами; некоторые ретропозоны не имеют ДКП.
- Существуют элементы, порожденные событиями транспозиции, опосредованными РНК, но при этом не кодирующие ферментов, катализирующих транспозицию.
- Транспозоны и ретропозоны составляют почти половину генома человека.

Принадлежность к ретропозонам определяется свойством мобильного элемента использовать обратную транскрипцию РНК в ДНК как составную часть механизма

В геномах эукариот содержатся ретропозоны трех типов			
	Вирусное надсемейство	LINES	Невирусное надсемейство
Типичные представители	Ty (*S. cerevisiae*) copia (*D. melanogaster*)	L1 (человек) B1, B2 ID, B4 (мышь)	SINES (млекопитающие), псевдогены транскриптов ДНК-полимеразы III
Края	Длинные концевые повторы	Нет повторов	Нет повторов
Повторы мишени	4–6 пн	7–21 пн	7–21 пн
Кодируемые активности	Обратная транскриптаза и/или интеграза	Обратная транскриптаза/ эндонуклеаза	Нет (во всяком случае, активностей, свойственных транспозонам)
Организация	Могут содержать интроны (в субгеномной мРНК они подлежат удалению)	Одна или две непрерывающихся открытых рамки считывания	Интронов нет

РИС. 22.16 Ретропозоны эукариот можно подразделить на три надсемейства. Представители двух надсемейств (вирусного и LINES) довольно близки к ретровирусам. Ретропозоны невирусного надсемейства являются вырожденными и не кодируют никаких собственных белков

транспозиции. Подробности разбиения ретропозонов на три систематические группы раскрывает РИС. 22.16.

- Члены надсемейства **вирусных ретропозонов** кодируют обратную транскриптазу и/или интегразу. Репродуцируются они, как и другие ретропозоны, подобно ретровирусам, от которых отличаются отсутствием независимой формы с инфекционным потенциалом. Самые изученные из них — *Ty* и *copia* у дрожжей и мух соответственно.
- Длинные рассеянные повторяющиеся последовательности (long interspersed repeated sequences, LINES) также обладают закодированной в них ревертазной активностью (и, возможно, могут считаться дальними родственниками вирусных ретропозонов), однако у них нет ДКП, и они пользуются другим механизмом затравления обратной транскрипции, чем ретровирусы. Они ведут свое происхождение от транскриптов РНК-полимеразы II. LINES, которые полностью функциональны и автономны в плане транспозиции, в геноме меньшинство; остальные несут мутации и способны совершить транспозицию только в результате энзиматической *транс*-активности автономного элемента.
- Членов надсемейства **невирусных ретропозонов** отличают особенности, позволяющие предположить их происхождение от последовательностей РНК. Все же относительно происхождения ДНК-копий РНК-матрицы остается только спекулировать. Они всегда неавтономны и, предположительно, подпали под транспозицию силами ферментов, закодированных в другом месте. Невирусные ретропозоны ведут свое происхождение от клеточных трнаскриптов и не кодируют белков, которые участвовали бы в транспозиции. Самые известные члены надсемейства называются SINES (short interspersed repeated sequences, короткие рассеянные повторяющиеся последовательности) и являются производными продуктов активности РНК-полимеразы III.

На РИС. 22.17 показано, как соотносятся между собой структуры и последовательности разных элементов, кодирующих обратную транскриптазу (ревертазу). Подобно ретровирусам, ретропозоны, содержащие ДКП, могут быть подразделены на систематические группы по числу независимых открытых рамок считывания для гомологов *gag*, *pol* и *int*, а также по порядку, в котором эти рамки расположены. Несмотря на небольшие различия в организации, их общей чертой является наличие собственных ревертазы и интегразы. Типичные LINES млекопитающих содержат две открытых рамки считывания; одна из которых кодирует белок, связывающий нуклеиновую кислоту, другая — активности ревертазы и эндонуклеазы.

ДКП-содержащие элементы — это и ретровирусы, и ретропозоны, утратившие способность генерировать инфекционные частицы. Геномы дрожжей и мух содержат элементы, соответственно *Ty* и *copia*, неспособные генерировать инфекционные частицы. Геном мыши содержит несколько активных эндогенных ретровирусов, способных генерировать частицы, распространяющие инфек-

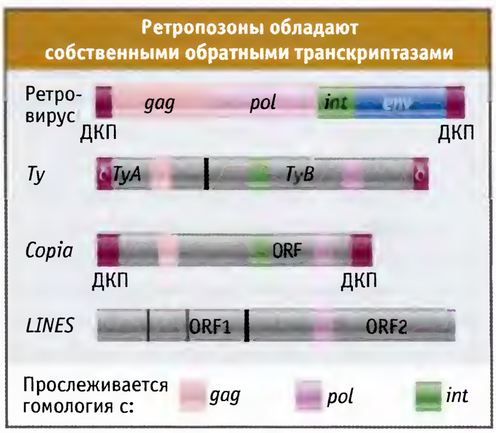


РИС. 22.17 Ретропозоны, близкородственные ретровирусам, имеют похожую с ними организацию. У ретропозонов LINES сходство с ретровирусами исчерпывается наличием собственной обратной транскриптазы

Ретровирусы и транспозоны составляют половину генома человека

Элемент	Организация	Длина (тпн)	В геноме человека: Число	Доля
Ретропозоны, близкие к ретровирусам	ДКП *gag* *pol* *(env)* ДКП	1–11	450 000	8%
LINES (автономные), например, L1	*ORF1* *(pol)* (A)$_n$	6–8	850 000	17%
SINES (неавтономные), например, Alu	(A)$_n$	<0,3	1 500 000	15%
ДНК-транспозоны	Транспозаза	2–3	300 000	3%

РИС. 22.18 Почти половина генома человека занята мобильными элементами четырех типов

цию горизонтально (то есть заражать потомство других клеток). В линии предков человека, напротив, почти все эндогенные ретровирусы утратили свою активность около 50 миллионов лет тому назад, и на сегодняшний день человеческий геном содержит, по большей части, неактивные останки эндогенных ретровирусов.

Элементы LINES и SINES оккупируют значительную часть животных геномов. Исходно идентифицированные как многочисленные, относительно короткие, похожие друг на друга последовательности в составе умеренно повторяющейся ДНК (описанной в разд. 4.6 «Эукариотические геномы состоят из уникальных и повторяющихся последовательностей»). LINES и SINES получили определение рассеянных последовательностей, или **рассеянных повторов**, из-за своей частой встречаемости и пространного распределения в геномах.

В растениях присутствует другой тип малых мобильных элементов, называемый MITE (миниатюрный мобильный элемент с инвертированными повторами, «клещ»). Элементы MITE оканчиваются инвертированными повторами, фланкированы 2 пн или 3 пн повторов мишени, не кодируют белков и имеют длину от 200 до 500 пн. Например, в геноме риса представлено, по меньшей мере, девять семейств MITE. Их нередко можно обнаружить во фланкирующих областях генов, кодирующих белки. Никакого отношения к SINES или LINES они не имеют.

Как уже говорилось, LINES и SINES составляют значительную часть повторяющейся ДНК в геномах животных. Во многих геномах высших эукариот этими элементами оккупировано ~50% тотальной ДНК. РИС. 22.18 подытоживает данные о различных типах транспозонов, составляющих почти половину генома человека. За исключением SINES, которые всегда нефункциональны, все типы транспозонов представлены как функциональными элементами, так и элементами, претерпевшими делеции, с нарушенной открытой рамкой считывания, кодирующей белок, необходимый для транспозиции. В геноме мыши эти типы транспозонов тоже присутствуют, и в сходных пропорциях.

Самое распространенное в геномах млекопитающих семейство LINES называют L1. Его типичный представитель имеет длину 6500 пн и оканчивается A-богатым трактом. Две открытых рамки считывания полноразмерного L1-элемента называют: ORF1 и ORF2. Число полноразмерных L1-элементов в геноме обычно невелико (~50), остальные — укороченные. Можно обнаружить и L1-транскрипты. LINES вариабельны по последовательности, причем степень вариабельности LINES между разными видами животных гораздо выше, чем внутри одного вида. Единственные LINES, которые были активными и у предков мыши, и у предков человека, — это L1. У мышей они, по-видимому, так и остались высокоактивными, в то время как у людей их активность в процессе эволюции сошла на нет.

В эволюционной родословной человека из всех SINES активными были только Alu-элементы. В геноме мыши есть аналог Alu (B1), а также другие SINES (B2, ID, B4), когда-то бывшие активными. Вероятно, Alu-элементы человека и B1-элементы мыши происходят от 7СЛ РНК (см. разд. 22.10 «Alu-семейство и его эквиваленты в геномах млекопитающих»). Другие мышиные SINES, по всей видимости, представляют собой производные обратных транскриптов тРНК. Транспозиция SINES происходит, скорее всего, в результате опознавания их в качестве субстрата ферментами активного L1-элемента.

22.10 Alu-семейство и его эквиваленты в геномах млекопитающих

Ключевые понятия:

- Большая часть повторяющейся ДНК в геномах млекопитающих представлена элементами, принадлежащими к одному семейству, организованными подобно транспозонам и ведущими свое происхождение от продуктов активности РНК-полимеразы III.

Основная масса SINES — это члены родственных семейств элементов в геномах разных видов млекопитающих. По своим небольшим размерам и высокой повторяемости они похожи на сателлитную ДНК, но в отличие от нее не образуют в геноме тандемные кластеры, а распределены в нем рассеянно. Как и в случае LINES, степень сходства элементов внутри семейства (в геномах животных одного вида) превышает степень сходства представителей семейств у разных видов.

В геноме человека значительная часть умеренно повторяющейся ДНК состоит из наборов последовательностей в 300 пн, которыми перемежается неповторяющаяся ДНК. В экспериментах по ренатурации, как минимум половина материала геномной ДНК, восстановившегося в двухцепочечной форме, подвержена расщеплению эндонуклеазой рестрикции AluI по единственному сайту, расположенному внутри элемента SINES, в 170 пн от его начала. Все последовательности, разрезаемые таким образом, относят к одному семейству, названному **Alu-семейством**, в честь способа его идентификации. В гаплоидном геноме насчитывается ~300 000 членов (элементов) этого семейства, что соответствует одному элементу на 6 тпн ДНК. Похожие семейства последовательностей есть у мыши (B1, представленное ~50 000 элементов), у китайского хомячка (Alu-эквивалентное семейство), у других млекопитающих.

Члены Alu-семейства чаще схожи, но не идентичны. По всей видимости, элемент-прототип Alu-семейства человека возник благодаря тандемной дупликации, сопровождавшейся вставкой 31 пн чужеродной последовательности в правую половину димера. Эти два повтора иногда называют левой и правой половинами последовательности Alu. Последовательности отдельных членов (элементов) Alu-семейства совпадают с консенсусом, в среднем,

на 87%. У мыши повторяющаяся единица B1 имеет длину 130 пн и соответствует мономеру Alu-повтора человека. Сходство между ними составляет 70–80%.

Последовательность Alu родственна 7СЛ РНК, компоненту сигнал-узнающей частицы (см. разд. 10.9 «SRP взаимодействует со своим рецептором»). 7СЛ РНК соответствует левой половинке последовательности Alu со вставкой в середине. Таким образом, 90 5′-концевых оснований 7СЛ РНК гомологичны левому краю левой половинки последовательности Alu, центральные 160 оснований 7СЛ РНК не имеют гомологии с последовательностью Alu, а 40 3′-концевых оснований гомологичны правому краю левой половинки последовательности Alu. 7СЛ РНК кодируется генами, которые активно транскрибирует РНК-полимераза III. Возможно, эти гены (или родственные им) и дали начало неактивным Alu-элементам.

Члены Alu-семейства, подобно транспозонам, фланкированы короткими прямыми повторами. Любопытно, что длина этих повторов индивидуальна для разных членов семейства. Alu-семейство происходит от транскриптов РНК-полимеразы III, поэтому отдельные его представители могут нести внутри себя активные промоторы.

Многообразие свойств и вездесущность Alu-семейства питает множество гипотез по поводу его функции. Несмотря на это, постичь его истинную роль пока не удалось.

По меньшей мере некоторые из членов Alu-семейства могут быть транскрибированы в независимые РНК. У китайского хомячка некоторые (но не все) члены Alu-эквивалентного семейства, по-видимому, транскрибируются *in vivo*. Единицы транскрипции такого рода можно обнаружить поблизости от других единиц транскрипции.

Как можно заметить по их присутствию во фракции длинной ядерной РНК, члены Alu-семейства могут входить в состав единиц транскрипции структурных генов. Присутствие нескольких копий последовательности Alu внутри одной молекулы РНК может придать ей выраженную вторичную структуру. В самом деле, львиная доля вторичной структуры ядерной РНК млекопитающих обусловлена присутствием в ней инвертированных повторов, принадлежащих последовательностям Alu.

22.11 Процессированные псевдогены ведут свое происхождение от субстратов транспозиции

Ключевые понятия:

- Процессированный псевдоген представляет собой производную последовательности мРНК, полученной в ходе обратной транскрипции.

Распознать последовательность, полученную в ходе обратной транскрипции мРНК и затем встроенную в геном, можно, сравнив эту последовательностью с геном, послужившим матрицей для синтеза ее мРНК. ДНК-копию мРНК, встроенную в геном клетки, называют **процессированным псевдогеном**, в отражение того обстоятельства, что эта модифицированная копия гена была процессирована, как РНК, и оттого неактивна. На РИС. 22.19 проведено сравнение процессированного псевдогена с геном-оригиналом и его мРНК. Рисунок сводит воедино все диагностически-значимые признаки; как правило, лишь часть их обнаруживается в каждом отдельно взятом примере процессированного псевдогена. Дать начало псевдогену может любой транскрипт, синтезированный РНК-полимеразой II. Существует множество примеров таких событий, в их числе — процессированные псевдогены глобинов, открытые первыми (см. разд. 3.11 «Вся ли генетическая информация содержится в ДНК?»).

Псевдоген может начинаться в точке, эквивалентной 5′-концу мРНК, чего можно было бы ожидать лишь в случае происхождения его ДНК от РНК. Некоторые псевдогены состоят из точно состыкованных последовательностей экзонов; насколько известно, никакого механизма опознания интронов в составе ДНК не существует, и такая особенность строения псевдогена говорит о его РНК-опосредованном происхождении. Псевдоген может оканчиваться коротким отрезком пар оснований А-Т, предположительно производным от поли(А)-хвоста РНК. Псевдоген фланкируют короткие прямые повторы, как если бы он появился в результате транспозиции. Хромосомная локализация процессированного псевдогена никак не связана с его предполагаемым местом происхождения.

Процессированные псевдогены не несут никакой информации, которая могла бы быть использована в поддержку события транспозиции (или хотя бы для предварительной обратной транскрипции РНК). Это указывает на то, что РНК (предшественник псевдогена) служила суб-



РИС. 22.19 Псевдогены возникают вследствие встраивания двухцепочечных продуктов обратной транскрипции мРНК в геном

стратом другой системе, кодируемой ретропозоном. Основной источник клеточной ревертазной активности — это элементы LINES. Вполне возможно, что они, помимо обеспечения собственной транспозиции, воздействуют и на последовательности SINES, и стимулируют возникновение процессированных псевдогенов.

22.12 LINES использует собственную эндонуклеазу, чтобы создать затравочный конец для синтеза ДНК

Ключевые понятия:

- У элементов LINES отсутствуют ДКП. Поэтому эти ретропозоны кодируют эндонуклеазу, которая производит надрезы для затравки обратной транскрипции.

В отличие от типичных ретровирусных элементов, члены семейства LINES и некоторые другие ретропозоны не фланкированы ДКП. Как происходит затравление обратной транскрипции в таких случаях? Обычная реакция, при которой тРНК-затравка залипает на ДКП (см. рис. 22.6) здесь неприменима. Открытые рамки считывания LINES не кодируют белковых доменов с активностями протеазы и интегразы, но обычно содержат последовательность, гомологичную ревертазе, а также кодируют собственную эндонуклеазу. В LINES L1 человека ORF1 соответствует ДНК-связывающему белку, а ORF2 — активностям ревертазы и эндонуклеазы. Для транспозиции необходимы оба продукта.

Как именно они задействованы, показано на РИС. 22.20. Сначала кодируемая ретропозоном LINES эн-

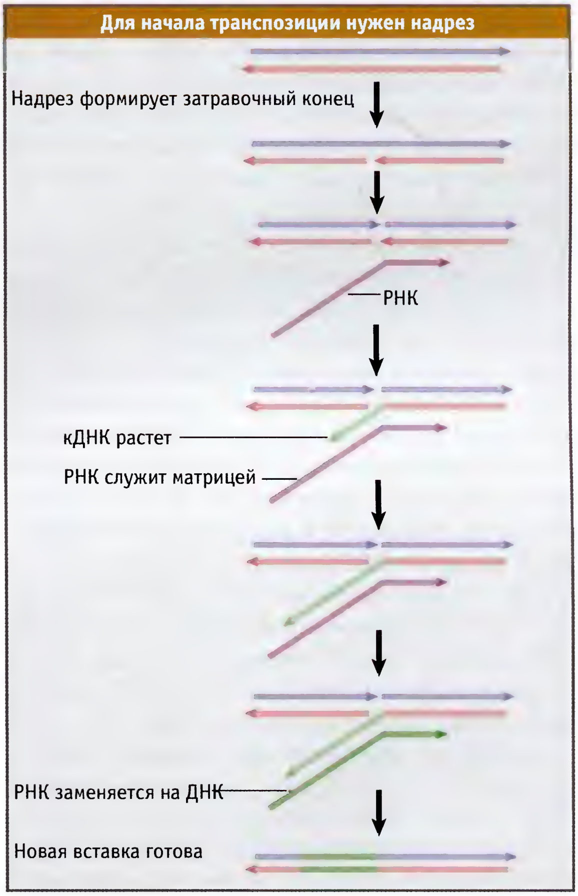


РИС. 22.20 Ретротранспозиция элементов, у которых отсутств ДКП, в принципе, возможна. Для ее начала необходимо поя ние надреза в сайте-мишени; из него берется свободный за вочный конец, необходимый для обратной транскрипции. В зультате в сайте-мишени (на матрице РНК) синтезируется кД Событие рекомбинации может превратить ее во вставку. 3′-ко цепей обозначены стрелками

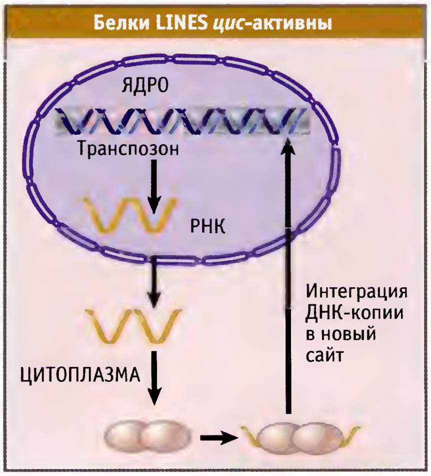


РИС. 22.21 Геномная копия LINES транскрибируется в РНК, которая транслируется в белки, образующие с ней комплекс. Этот комплекс направляется в ядро и вставляет в геном очередную ДНК-копию LINES

донуклеаза вносит надрез (ник) в сайт-мишень. Таким образом, формируется 3'-OH-конец, затравляющий синтез кДНК на матрице РНК. РНК-продукт элемента (субстрат транспозиции) ассоциирует с белком, связывающим ДНК по месту надреза. Второй надрез вскрывает оставшуюся цепь ДНК, после чего ДНК-РНК гибрид подшивается своим противоположным концом к другому флангу разреза (это может произойти позже, уже после превращения субстрата транспозиции в дуплекс ДНК, как и показано на рисунке). Похожий механизм используется некоторыми мобильными интронами (см. рис. 27.12).

Если элементы ведут свое происхождение от транскриптов РНК-полимеразы II, их геномные последовательности неизбежно неактивны: в них отсутствует промотор, находившийся выше точки старта транскрипции. Обычно такие элементы обладают признаками зрелого транскрипта, потому их и называют **процессированными псевдогенами**.

Одна из причин, по которым LINES столь эффективны в распространении по геному своих процессированных псевдогенов, заключается в самом механизме этого распространения. Когда мРНК LINES, оказавшись в цитоплазме, транслируется, ее белковые продукты выказывают *цис*-предпочтение к связыванию именно той мРНК, с которой они были транслированы. На РИС. 22.21 показано, как после такого связывания рибонуклеопротеиновый комплекс движется в ядро, где белки вставляют в геном ДНК-копию мРНК. Нередко обратная транскрипция не доводится до конца, так что копия получается неактивной. Однако потенциально вставка может оказаться активной, ведь субстратом транспозиции служил транскрипт *активного* элемента (хоть и без промотора).

Белковые продукты эукариотических ДНК-транспозонов, чья транспозиция не поддерживается РНК-субстратом, напротив, не располагают возможностью отличить полноразмерный транспозон от неактивного (делетированного). Когда после синтеза в цитоплазме они оказываются импортированными в ядро, то используют в качестве субстрата любой элемент, маркированный инвертированными концевыми повторами с подходящей структурой (см. РИС. 22.22). Это во много раз снижает шанс того, что они подействуют на полноразмерный, а не делетированный, элемент. Следствием является накопление неактивных элементов. Постепенно семейство вымирает, поскольку вероятность встречи транспозазы с полностью функциональным транспозоном в качестве субстрата становится чрезвычайно мала.

Идет ли транспозиция в геномах современных млекопитающих, или же нашему наблюдению предстают лишь следы древних мобильных систем? Это зависит от конкретного вида млекопитающих. В геноме человека активных транспозонов мало, зато в геноме мыши насчитывается несколько их разновидностей. Поэтому спонтанные мутации по причине внедрения LINES случаются у мыши с частотой ~3%, а у человека — всего ~0,1%. Активных элементов LINES в человеческом геноме может быть от ~10 до ~50. Некоторые заболевания человека могут развиться из-за транспозиций L1 внутрь генов, другие — из-за неравного кроссинговера при участии повторяющихся копий L1. Исследования, проведенные на модельной системе, где транспозиции LINES имели место в клетках культуры тканей, показали, что сопутствующий ущерб от внедрения LINES может заключаться в хромосомных перестройках и делециях — мощных механизмах генетических изменений. В геномах (клеток зародышевого пути) предков человека ни ДНК-транспозоны, ни ретровирусные транспозоны не были активными уже 40–50 миллионов лет, в то время как у современных мышей обнаруживается по нескольку экземпляров и тех, и других.

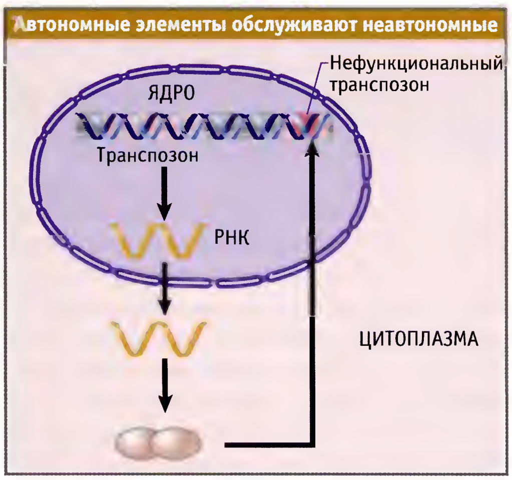


РИС. 22.22 Транспозон транскрибируется в РНК, которая транслируется в белки. Эти белки независимо направляются в ядро, где действуют на любую пару инвертированных повторов (с той же нуклеотидной последовательностью, что и у сходного транспозона), которую им удастся распознать

Заметим, что для закрепления транспозиции в геномах потомков она должна произойти в линии клеток зародышевого пути. Транспозиция, которая произошла в соматической клетке, недействительна в отношении последующих поколений организмов.

22.13 Резюме

Обратная транскрипция — это общий механизм размножения ретровирусов и самоподдержания ретропозонов. Циклы самовоспроизведения этих двух элементов эукариотического генома похожи, хотя умозрительно генетический материал ретровирусов привычнее отождествлять со свободной вирусной формой (РНК), ретропозоны же — рассматривать как элемент генома клетки-хозяина (двухцепочечной ДНК).

Ретровирусы имеют одноцепочечные РНК-геномы, которые для репликации переходят в промежуточную форму двухцепочечной ДНК. Каждый отдельно взятый ретровирус содержит две копии РНК-генома. В нем содержатся гены *gag*, *pol* и *env*, транслируемые в полипротеины, каждый из которых претерпевает расщепление до небольших функциональных белков. Компоненты Gag и Env задействованы в упаковке РНК и создании вириона, компоненты Pol задействованы в синтезе нуклеиновых кислот.

Главным из компонентов Pol является обратная транскриптаза (ревертаза). Этот фермент отвечает за синтез (–)-цепи — ДНК-копии вирусной РНК, или (+)-цепи. ДНК-продукт получается длиннее, чем РНК-матрица. Меняя матрицу с одной цепи РНК на другую, ревертаза прикрепляет копию 5′-концевой последовательности РНК к 5′-концу (–)-цепи ДНК (то есть, в конечном счете, к 3′-концу (+)-цепи ДНК). Похожая смена матрицы происходит в ходе синтеза (+)-цепи ДНК на матрице (–)-цепи ДНК; при этом 3′-концевая последовательность РНК оказывается на 5′-конце (+)-цепи ДНК. Так возникают характерные ДКП (длинные концевые повторы) ДНК ретровируса. Линейная двухцепочечная ДНК ретровируса вставляется в геном хозяина ферментом интегразой. Транскрипция интегрированной ДНК с промотора в левом ДКП генерирует последующие копии последовательности РНК ретровируса.

Смена матрицы в процессе синтеза нуклеиновой кислоты делает возможной рекомбинацию по механизму выбора копии. В течение инфекционного цикла вирус может обменять часть своей обычной последовательности на фрагмент клеточного генома. Получившийся вирус репликационно-дефективен, но может размножиться в условиях совместной инфекции с вирусом-помощником. Многие из таких дефектных вирусов заполучили в свой состав РНК-версию (*v-onc*) клеточного гена (*c-onc*). Последовательность *onc* может принадлежать любому из многочисленных генов, чья экспрессия в форме *v-onc* приводит к трансформации клетки в опухолеродный фенотип.

Событие интеграции ДНК вируса в клеточный геном порождает прямые повторы последовательности сайта-мишени (подобно тому, как это делают транспозоны, мобилизуемые через промежуточную форму ДНК). Таким образом, внедрившийся провирус несет прямые ДКП и фланкирован короткими прямыми повторами ДНК-мишени. Геномы млекопитающих и птиц содержат эндогенные (неактивные) провирусы с такой структурой. Другие элементы с такой же организацией обнаруживаются в самых разных эукариотических геномах, в особенности — у *S. cerevisiae* и *D. melanogaster*. Элементы *Ty* и *copia* в геномах дрожжей и мух, соответственно, несут кодирующие последовательности, гомологичные ревертазе, и мобилизуются через промежуточную форму РНК. Они могут генерировать частицы, напоминающие вирусные, которые, однако, не обладают способностью инфицировать другие клетки. Последовательности LINES в геномах млекопитающих ушли от ретровирусов еще дальше, сохранив, впрочем, степень сходства, говорящую об общности происхождения с ретровирусами. Чтобы инициировать обратную транскрипцию, LINES используют альтернативный способ прайминга: эндонуклеазная активность, ассоциированная с обратной транскриптазой, вносит в ДНК-мишень надрез (ник), из которого заимствует 3′-ОН-конец, затравляющий синтез ДНК на матрице LINES-РНК. В сопоставлении с элементами с ДНК-опосредованной мобильностью, частота транспозиций у LINES повышенная. Это результат того, что их белковые продукты *цис*-активны: они отличаются сродством к той мРНК, с которой были транслированы, и ассоциируют с ней, образуя нуклеопротеиновый комплекс, который транспортируется в ядро.

Представители другого надсемейства ретропозонов, обладая приметами транспозиции через РНК, не несут кодирующих последовательностей (по крайней мере, таких, что имели бы отношение к ретровирусным функциям). Они могли возникнуть как пассажиры события транспозиции на ретровирусный манер, где РНК послужила обратной транскриптазе субстратом. Так возникли процессированные псевдогены. Особая разновидность элементов, порожденных процессингом в сочетании с обратной транскрипцией и интеграцией, — это SINES млекопитающих; к ним относят Alu-семейство в геноме человека. Некоторые малые ядерные РНК (среди которых 7SL snРНК, компонент сигнал-узнающей частицы) родственны членам этого семейства.

Литература

22.2 Жизненный цикл ретровируса включает в себя стадии, подобные транспозиции

Обзор

Varmus, H. E. and Brown, P. O. (1989). Retroviruses. In Howe, M. M. and Berg, D. E., eds., *Mobile DNA*. Washington, DC: American Society for Microbiology, pp. 53–108.

Исследование

Baltimore, D. (1970). PHK-dependent DNA polymerase in virions of PHK tumor viruses. *Nature* 226, 1209–1211.

Temin, H. M. and Mizutani, S. (1970). PHK-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* 226, 1211–1213.

22.4 ДНК ретровируса является продуктом обратной транскрипции

Обзор

Katz, R. A. and Skalka, A. M. (1994). The retroviral enzymes. *Annu. Rev. Biochem.* 63, 133–173.

Lai, M. M. C. (1992). PHK recombination in animal and plant viruses. *Microbiol. Rev.* 56, 61–79.

Negroni, M. and Buc, H. (2001). Mechanisms of retroviral recombination. *Annu. Rev. Genet.* 35, 275–302.

Исследование

Hu, W. S. and Temin, H. M. (1990). Retroviral recombination and reverse transcription. *Science* 250, 1227–1233.

Negroni, M. and Buc, H. (2000). Copy-choice recombination by reverse transcriptases: reshuffling of genetic markers mediated by PHK chaperones. *Proc. Natl. Acad. Sci. USA* 97, 6385–6390.

22.5 Вирусная ДНК интегрируется в хромосому

Обзор

Goff, S. P. (1992). Genetics of retroviral integration. *Annu. Rev. Genet.* 26, 527–544.

Исследование

Craigie, R., Fujiwara, T., and Bushman, F. (1990). The IN protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration *in vitro*. *Cell* 62, 829–837.

22.7 Ty-элементы дрожжей аналогичны ретровирусам

Исследование

Boeke, J. D. et al. (1985). Ty elements transpose through an PHK intermediate. *Cell* 40, 491–500.

22.8 В числе мобильных элементов *Drosophila melanogaster* есть ретропозоны

Исследование

Mount, S. M. and Rubin, G. M. (1985). Complete nucleotide sequence of the *Drosophila* transposable element copia: homology between copia and retroviral proteins. *Mol. Cell Biol.* 5, 1630–1638.

22.9 Систематика ретропозонов

Обзор

Deininger, P. L. (1989). SINEs: short interspersed repeated DNA elements in higher eukaryotes. In Howe, M. M. and Berg, D. E., eds., *Mobile DNA.* Washington, DC: American Society for Microbiology, pp. 619–636.

Hutchison, C. L. A., Hardies, S. C., Loer, D. D., Shehee, W. R., and Edgell, M. H. (1989). LINES and related retroposons: Long interspersed repeated sequences in the eukaryotic genome. In Howe, M. M. and Berg, D. E., eds., *Mobile DNA.* Washington, DC: American Society for Microbiology, pp. 593–617.

Ostertag, E. M. and Kazazian, H. H. (2001). Biology of mammalian L1 retrotransposons. *Annu. Rev. Genet.* 35, 501–538.

Weiner, A. M., Deininger, P. L., and Efstratiadis, A. (1986). Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. *Annu. Rev. Biochem.* 55, 631–661.

Исследование

Bureau, T. E., Ronald, P. C., and Wessler, S. R. (1996). A computer-based systematic survey reveals the predominance of small inverted-repeat elements in wild-type rice genes. *Proc. Natl. Acad. Sci. USA* 93, 8524–8529.

Bureau, T. E. and Wessler, S. R. (1992). Tourist: a large family of small inverted repeat elements frequently associated with maize genes. *Plant Cell* 4, 1283–1294.

Loeb, D. D. et al. (1986). The sequence of a large L1Md element reveals a tandemly repeated 5′ end and several features found in retrotransposons. *Mol. Cell Biol.* 6, 168–182.

Sachidanandam, R. et al. (2001). A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. The IntePHKtional SNP Map Working Group. *Nature* 409, 928–933.

Waterston et al. (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature* 420, 520–562.

22.12 LINES использует собственную эндонуклеазу, чтобы создать затравочный конец для синтеза ДНК

Обзор

Ostertag, E. M. and Kazazian, H. H. (2001). Biology of mammalian L1 retrotransposons. *Annu. Rev. Genet.* 35, 501–538.

Исследование

Feng, Q., Moran, J. V., Kazazian, H. H., and Boeke, J. D. (1996). Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* 87, 905–916.

Gilbert, N., Lutz-Prigge, S., and Moran, J. V. (2002). Genomic deletions created upon LINE-1 retrotransposition. *Cell* 110, 315–325.

Lauermann, V. and Boeke, J. D. (1994). The primer tPHK sequence is not inherited during Ty1 retrotransposition. *Proc. Natl. Acad. Sci. USA* 91, 9847–9851.

Luan, D. D. et al. (1993). Reverse transcription of R2Bm PHK is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* 72, 595–605.

Moran, J. V., Holmes, S. E., Naas, T. P., DeBerardinis, R. J., Boeke, J. D., and Kazazian, H. H. (1996). High frequency retrotransposition in cultured mammalian cells. *Cell* 87, 917–927.

Symer, D. E., Connelly, C., Szak, S. T., Caputo, E. M., Cost, G. J., Parmigiani, G., and Boeke, J. D. (2002). Human l1 retrotransposition is associated with genetic instability *in vitro*. *Cell* 110, 327–338.

Как формируется многообразие антител

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

23.10 Белки RAG катализируют разрыв и воссоединение ДНК

- Для реакции расщепления необходимо и достаточно активности белков RAG.
- RAG1 распознает консенсусную последовательность-нонамер для рекомбинации. RAG2 связывается с RAG1 и расщепляет ДНК в районе гептамера.
- Реакция напоминает события, происходящие при разрешении цепей в результате работы топоизомеразы при транспозиции.
- На конца, несущем кодирующую последовательность, образуется промежуточный продукт в виде шпильки; раскрывание этой шпильки приводит к внедрению дополнительных оснований (P-нуклеотидов) в образующийся ген.
- Дезоксинуклеозид-трансфераза вводит добавочные N-нуклеотиды в кодирующий конец.
- Кодон в месте объединения V-(D)-J сегментов исключительно вариабелен и кодирует аминокислоту 96 в районе антиген-связывающего участка.
- Двухцепочечные разрывы в месте объединения кодирующих последовательностей репарируются той же системой, что действует при соединении негомологичных концов поврежденной ДНК.
- Энхансер в С-гене активирует промотор V-гена после рекомбинации (в случае образования интактного гена иммунолглобулина).

23.11 Процессинг РНК может изменять раннюю экспрессию тяжелой цепи

- Все лимфоциты начинают экспрессировать иммуноглобулины с мембраносвязанной формы IgM.
- Когда В-клетки дифференцируются, мембраносвязанная форма сменяется секреторной за счет сплайсинга РНК.

23.12 Рекомбинация ДНК вызывает переключение классов иммуноглобулинов

- Иммуноглобулины подразделяют на пять классов в соответствии с типом константного участка тяжелой цепи.
- Переключение классов при смене C_H-участка происходит путем рекомбинации между S-районами с делецией всей последовательности между старым и новым C_H-участками.
- Может происходить множество успешных рекомбинаций для переключения классов.

23.13 Переключение происходит путем необычной реакции рекомбинации

- Переключение происходит с помощью двухцепочечного разрыва и соединения негомологичных частей.
- Характерной особенностью района переключения является наличие инвертированных повторов.
- Для переключения необходима активация промоторов выше сайта переключения.

23.14 Соматические мутации создают дополнительное многообразие антител у мышей и человека

- Последовательность V-районов активных генов иммуноглобулинов отличается от этого участка в клетках зародышевого типа из-за соматических мутаций.
- Соматические мутации представляют собой замену конкретных нуклеотидов.
- Сайты мутаций концентрируются в области антиген-связывающего участка.
- Внесение соматических мутаций зависит от работы энхансера, который активирует транскрипцию Ig-локуса.

23.15 Цитидиндезаминаза и урацилгликозилаза вызывают соматический мутагенез

- Цитидин-дезаминаза необходима как для соматического мутагенеза, так и для переключения классов иммуноглобулинов.
- Активность урацил-ДНК-гликозилазы влияет на структуру соматических мутаций.
- Последовательное действие этих ферментов может вызвать гипермутацию.

23.16 Иммуноглобулины птиц образуются из псевдогенов

- Ген иммуноглобулина у цыпленка формируется путем копирования последовательности одного из 25 псевдогенов в единственный активный локус гена.

23.17 Быстрый вторичный ответ возможен благодаря В-клеткам памяти

- Первичный ответ на антиген формируется благодаря В-клеткам, погибающие после ответа.
- Возникают неактивные В-клетки памяти, имеющие специфичность к тому же самому антигену.
- Повторная встреча с антигеном запускает вторичный ответ, и клетки памяти быстро активируются.

23.18 Т-клеточные рецепторы схожи с иммуноглобулинами

- Для образования одного из двух типов Т-клеточного рецептора Т-клетки используют аналогичный В-клеточному механизм V-(D)J-C-соединения.
- Т-рецепторы $\alpha\beta$ обнаруживаются на поверхности более 95% Т-лимфоцитов; Т-рецепторы $\gamma\delta$ — на поверхности менее 5% Т-лимфоцитов.

23.19 Т-клеточные рецепторы работают во взаимодействии с МНС

- Т-рецепторы распознают короткий пептид, расположенный в бороздке белка главного комплекса гистосовместимости на поверхности презентирующей клетки.

23.20 Локус главного комплекса гистосовместимости кодирует множество генов иммунной системы

- Локус МНС кодирует белки класса I и II, а также другие белки иммунной системы.
- Белки класса I — антигены трансплантации, отвечающие за распознавание «своих» и «чужих» тканей.
- Белки МНС класса I активны в виде гетеродимера с β_2-микроглобулином.
- Белки класса II работают при взаимодействии между Т-клетками.
- Белок МНС класса II представляет собой гетеродимер из α- и β-цепей.

23.21 Врожденный иммунитет в своей работе использует консервативные сигнальные пути

- Врожденный иммунитет активируется рецепторами (PAMP), распознающими определенные высококонсервативные мотивы у бактерий и других инфекционных агентов.
- Toll-подобные рецепторы широко используются для активации этого пути.
- Система врожденного иммунитета высококонсервативна у беспозвоночных и позвоночных, аналогичная система обнаружена и у растений.

23.22 Резюме

23.1 Введение

Все мы помним аксиому генетики, гласящую, что генотип образуется в зиготе путем комбинации генетического материала сперматозоида и яйцеклетки и затем наследуется всеми клетками взрослого организма. Чтобы объяснить различия фенотипов соматических клеток, мы исследуем различия в контроле экспрессии генов, но не учитываем возможность изменения состава ДНК.

Тем не менее существуют особые случаи реорганизации определенных последовательностей ДНК, что используется для регуляции экспрессии генов либо для создания новых генов. Иммунная система представляет собой удивительный пример, когда последовательность генома изменяется, а рекомбинация создает новые активные гены в лимфоцитах. Другие известные примеры — это замена одной последовательности ДНК на другую для изменения типа спаривания у дрожжей и для образования новых поверхностных антигенов у трипаносом.

Иммунный ответ у позвоночных представляет собой защитную систему, которая способна отличать чужеродные белки от белков собственного организма. Чужеродный материал (или часть такого материала) распознается как **антиген**. Обычно антиген — это белок (или частица, присоединенная к белку), который попал в кровоток животного, например белок оболочки вируса. Встреча с антигеном запускает иммунный ответ, целью которого является *специфичное распознавание и уничтожение антигена.*

За иммунные реакции отвечают белые кровяные тельца — В- и Т-лимфоциты, а также макрофаги. Лимфоциты получили свои названия по тем тканям, в которых они образуются. У млекопитающих **В-клетки** созревают в костном мозге, в то время как **Т-клетки** созревают в тимусе. *Для участия в иммунном ответе в каждом из этих классов лимфоцитов происходит реорганизация ДНК.*

У иммунной системы есть множество способов уничтожения антигена, однако удобно разделить их на два основных типа. Какой из этих типов будет развивать иммунная система при встрече с чужеродной структурой, частично зависит от природы антигена. Ответ определяется тем, какие клетки в основном задействованы: Т- или В-клетки.

Гуморальный ответ зависит от работы В-клеток. Клетки секретируют антитела, представляющие собой белки-**иммуноглобулины**. *Производство **антитела**, специфичного к чужеродной молекуле, — это главный процесс, отвечающий за распознавание антигена.* При распознавании антитело связывается с небольшим районом или структурой антигена.

Функция антител схематично представлена на РИС. 23.1. Чужеродный материал, циркулирующий в крови, — к примеру, токсин или патогенная бактерия, — на поверхности несет презентированные антигены. Антиген(ы) распознаются антителами, которые образуют комплекс антиген-антитело. Этот комплекс затем привлекает внимание других компонентов иммунной системы.

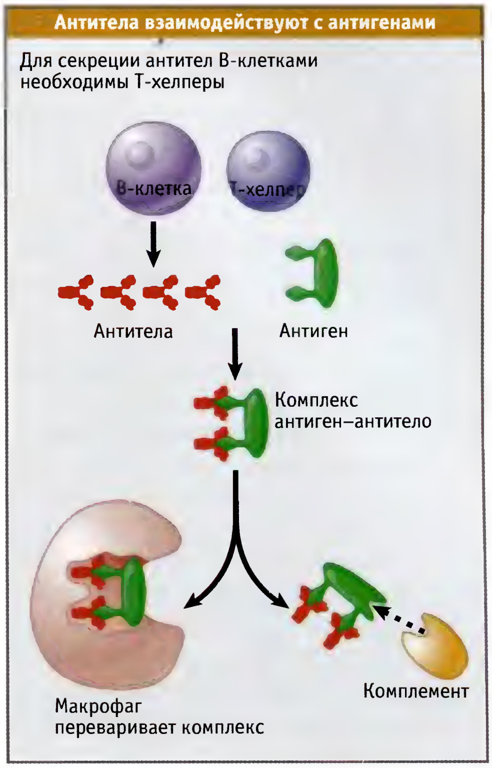


РИС. 23.1 Гуморальный иммунный ответ осуществляется путем связывания свободных антител с антигеном с формированием комплексов антиген-антитело, которые удаляются макрофагами или напрямую атакуются белками комплемента

Гуморальный ответ зависит от других компонентов иммунной системы. Во-первых, В-клетки должны получить сигналы от особых Т-клеток, чтобы начать секретировать антитела. Эти Т-клетки называют **Т-хелперами**, поскольку они оказывают помощь В-клеткам. Во-вторых, образование комплекса антиген-антитело инициирует процесс уничтожения антигена. Основной путь уничтожения — действие системы **комплемента**, компонента иммунной системы, чье название отражает его способность дополнять, «комплементировать» действие антитела. Система комплемента включает в себя около 20 белков, которые осуществляют свои функции путем каскада протеолитических реакций. В случае, когда антиген является частью клетки — например, инфицирующей бактерии, — функционирование системы комплемента приводит к лизису этой чужеродной клетки. Результат действия системы комплемента также служит для привлечения внимания макрофагов, удаляющих чужеродные клетки и продукты их жизнедеятельности. С другой стороны, комплекс антиген-антитело может уничтожаться непосредственно макрофагами (клетками-мусорщиками).

Особый класс Т-лимфоцитов — **цитотоксические Т-клетки** (Т-киллеры) — выполняет **клеточно-опосредованный ответ**. Основная функция

Т-клеток по распознаванию целевого антигена схематично представлена на РИС. 23.2. Клеточно-опосредованный ответ обычно вызывается внутриклеточными паразитами, например вирусами, инфицирующими клетки организма. В результате вирусной инфекции фрагменты чужеродных (вирусных) антигенов экспонируются на поверхности клетки. Эти фрагменты распознаются **рецепторами Т-клеток (TCR)**, являющимися Т-клеточными эквивалентами антител, продуцируемых В-клетками.

Отличительной чертой процесса распознавания является то, что *антиген должен быть* представлен *специальным белком, являющимся членом* **главного комплекса гистосовместимости (МНС)**. Белок МНС на поверхности несет бороздку, которая связывает белковый фрагмент чужеродного антигена. Рецептор Т-клеток распознает пептидный фрагмент в комбинации с белком МНС. Каждый организм имеет свой характерный набор белков МНС. Они важны в реакциях отторжения трансплантата; участок ткани от одного человека при пересадке другому человеку отторгается именно ввиду различий в белках МНС у донора и реципиента, и это имеет большое значение

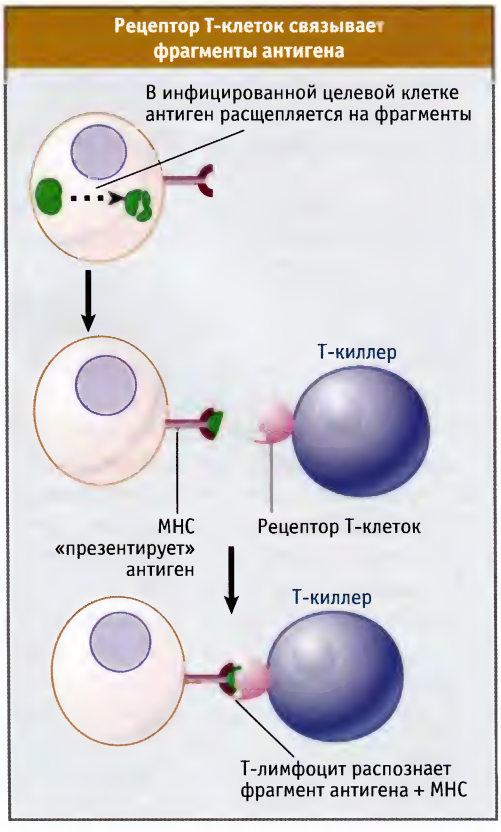


РИС. 23.2 При клеточно-опосредованном иммунном ответе киллерные Т-клетки используют рецепторы Т-клеток для распознавания фрагмента чужеродного антигена, презентированного на поверхности целевой клетки белком МНС

в медицине. Выполнение условия, что Т-лимфоциты распознают чужеродный антиген и белок МНС, гарантирует, что клеточно-опосредованный ответ действует только на клетки хозяина, инфицированные чужеродным антигеном. (Разделение белков МНС на класс I и II мы обсудим позже, в разд. 23.20 «Локус главного комплекса гистосовместимости кодирует множество генов иммунной системы».)

Целью любого иммунного ответа является атака чужеродного агента. Распознавание этого агента — прерогатива В-клеточных иммуноглобулинов и рецепторов Т-клеток. Ключевой момент в их функции — это способность отличать «своих» и «чужих». Белки и клетки собственного организма *никогда* не должны быть атакованы. А чужеродные агенты должны быть *полностью уничтожены*. Свойство не атаковать свои клетки называют **толерантностью**. Утрата этой способности приводит к **аутоиммунному заболеванию**, при котором иммунная система атакует собственный организм, что часто имеет разрушительные последствия.

Каков же механизм предотвращения реакции лимфоцитов на белки своего организма? По-видимому, толерантность возникает на ранних стадиях развития лимфоцитов, когда В-клетки и Т-клетки, распознающие «свои» антигены, уничтожаются. Этот процесс называют **клональной делецией**. В дополнение к этой негативной селекции есть и позитивная, с отбором Т-клеток, несущих определенные наборы рецепторов Т-клеток.

Из-за толерантности может оказаться очень трудно получить антитела к белкам, близкородственным белкам собственного организма. С практической точки зрения бывает сложно трудно использовать мышь или кролика, к примеру, для получения антител к белкам человека, являющимся высококонсервативными в эволюции млекопитающих. Толерантность мыши и кролика к собственным белкам может в этом случае распространяться и на человеческие белки.

Белки в каждой из трех групп, необходимых для развития иммунного ответа — иммуноглобулины, Т-клеточные рецепторы и белки МНС, — очень разнообразны. Исследуя большое число людей, можно найти множество вариантов этих белков. Каждый белок кодируется большим семейством генов; в случае антител и рецепторов Т-клеток многообразие увеличивается реорганизацией ДНК в соответствующих лимфоцитах.

Иммуноглобулины и рецепторы Т-клеток — прямые аналоги, но каждый образуется своим типом лимфоцитов. Эти белки родственны по структуре, а их гены сходным образом организованы. Источники многообразия их почти одинаковы. Белки МНС также имеют некоторые общие черты с антителами, как и другие лимфоцит-специфичные белки. В отношении организации генов иммунной системы мы имеем дело с целым рядом родственных семейств генов, то есть с **суперсемейством**, которое могло образоваться из общего предшественника, определявшего примитивный иммунный ответ.

23.2 Численность лимфоцитов, отвечающих на определенные антигены, увеличивается благодаря клональной селекции

Ключевые понятия:

- Каждый B-лимфоцит продуцирует иммуноглобулин одного вида, а каждый T-лимфоцит продуцирует рецептор T-клеток одного вида.
- Существует огромное разнообразие иммуноглобулинов и рецепторов T-клеток.
- Связывание антигена с иммуноглобулином или рецептором T-клеток приводит к размножению конкретного клона клеток.

Термин «иммунный ответ» получил такое название из-за одного из главных его признаков. После встречи организма с антигеном он становится *иммунным* к новым инфекциям. До встречи с конкретным антигеном у организма нет возможности предотвратить его токсические эффекты. Эта способность приобретается во время иммунного ответа. После победы над инфекцией организм сохраняет способность быстро отвечать в случае повторной инфекции.

Эти особенности объясняются теорией **клональной селекции**, показанной на РИС. 23.3. Пул лимфоцитов включает в себя B-клетки и T-клетки, несущие огромное разнообразие иммуноглобулинов или T-клеточных рецепторов. *Каждый B-лимфоцит производит один тип иммуноглобулинов, способный распознавать только один конкретный антиген; сходным образом, каждый T-лимфоцит производит только один специфический T-клеточный рецептор.*

Незрелые нестимулированные B- и T-клетки морфологически неотличимы друг от друга. При встрече с антигеном B-клетка, образующая антитело, способное связываться с ним, или T-клетка, несущая рецептор, способный его распознавать, получают стимул к делению, вероятно, путем обратной связи с поверхности клетки, где происходит реакция антиген-антитело/рецептор. Стимулированные клетки развиваются в зрелые B- или T-лимфоциты, что включает в себя морфологические изменения, включающие, к примеру, увеличение в размерах (особенно это касается B-клеток).

Начавшуюся экспансию специфических B- или T-клеток при первой встрече с антигеном называют **первичным иммунным ответом**. Образуется большое число B- и T-лимфоцитов, специфичных к целевому антигену. Каждая популяция представляет собой клон исходной реагирующей клетки. B-клетки секретируют антитела в большом количестве, и они могут даже преобладать во всей популяции антител.

После успешного первичного иммунного ответа организм сохраняет B- и T-клетки, производящие соответствующие антитела или рецептор. Такие **клетки памяти** находятся в промежуточном состоянии между зрелыми и незрелыми клетками. Они не имеют всех признаков зрелых клеток, однако такие клетки долгоживущие и могут быстро превратиться в зрелые клетки. Присутствие таких клеток в организме позволяет быстро развить **вторичный иммунный ответ**, если животное снова встретится с тем же самым антигеном.

Пул незрелых лимфоцитов у млекопитающих включает в себя примерно 10^{12} клеток. В нем содержатся некоторые лимфоциты с уникальными особенностями (поскольку соответствующие антигены еще ни разу не встречались), в то время как некоторые другие представлены в количестве до 10^6 клеток (поскольку число клеток, отвечающих на антиген, увеличивается благодаря клональной селекции).

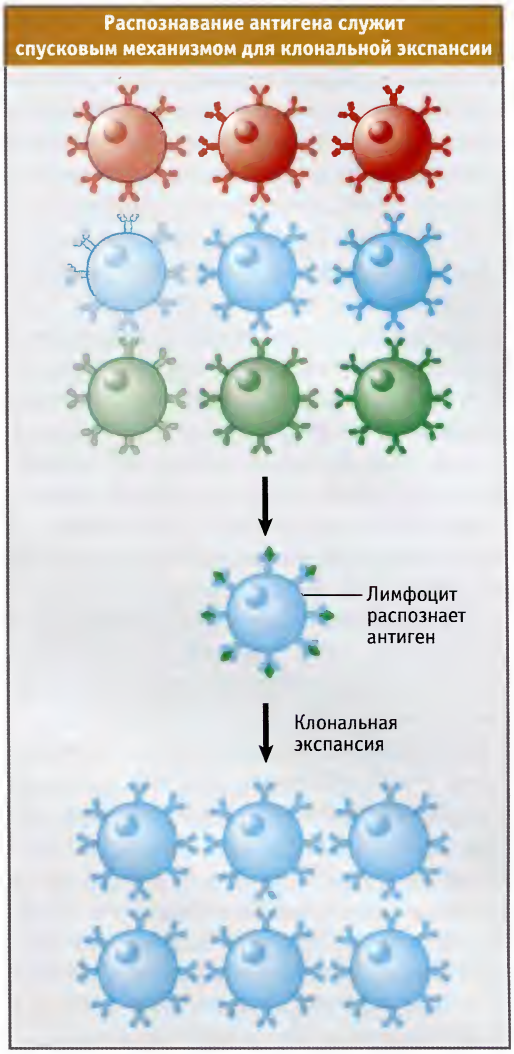


РИС. 23.3 Пул незрелых лимфоцитов включает в себя B-клетки и T-клетки, образующие антитела и рецепторы с разнообразными специфичностями. Реакция с антигеном приводит к клональной экспансии лимфоцита, образующего антитела (B-лимфоцит) или рецептор (T-лифоцит), распознающие данный антиген

Какие особенности распознаются в антигене? Антигены обычно являются макромолекулами. Небольшие молекулы могут иметь антигенные детерминанты и могут распознаваться антителами, хотя обычно они неэффективны в стимуляции иммунного ответа ввиду своего малого размера. Однако они могут вызывать иммунный ответ, если конъюгируют с большей несущей молекулой (обычно это белок). Небольшую молекулу, способную вызывать иммунный ответ по такому принципу, называют **гаптен**.

Лишь небольшая часть поверхности макромолекулярного антигена распознается одним конкретным антителом. Сайт связывания состоит всего из пяти-шести аминокислотных остатков. Разумеется, каждый белок может иметь больше одного сайта для связывания, в этом случае он распознается антителами со специфичностями к разным участкам. Область антигена, вызывающую ответ, называют **антигенной детерминантой,** или **эпитопом**. Если антиген содержит несколько эпитопов, некоторые из них могут быть более эффективны в вызывании иммунного ответа; фактически, они могут быть настолько эффективны, что будут доминировать в иммунном ответе.

Как лимфоциты находят целевые антигены и где происходит их созревание? Лимфоциты — странствующие клетки. Они развиваются из незрелых стволовых клеток, расположенных в костном мозге. Затем мигрируют в периферические лимфоидные ткани (селезенку, лимфатические узлы) либо напрямую с током крови (если это В-клетки), либо через тимус (где они становятся Т-клетками). Лимфоциты рециркулируют между кровью и лимфой; этот процесс гарантирует, что антиген будет экспонирован лимфоцитам со всеми возможными специфичностями. Когда лимфоцит встречает антиген, связывающий его антитело или рецептор, клональная экспансия начинает иммунный ответ.

23.3 Гены иммуноглобулинов компонуются из составных частей в лимфоцитах

Ключевые понятия:

- Иммуноглобулин — это тетрамер, состоящий из двух легких и двух тяжелых цепей.
- Легкие цепи формируют семейства лямбда и каппа; тяжелые цепи относятся к единственному семейству.
- В каждой цепи присутствует N-концевой вариабельный район (V) и С-концевой константный район (С).
- V-домен распознает антиген, а С-домен обеспечивает ответ эффекторной части.
- V-домены и С-домены кодируются отдельно сегментами V-гена и С-гена, соответственно.
- Ген, кодирующий иммуноглобулиновую цепь, образуется в ходе соматической рекомбинации путем объединения сегментов V- и С-генов.

Отличительной чертой иммунного ответа является способность животного производить соответствующее антитело при встрече с новым антигеном. Каким же образом организм может производить белковые антитела, каждое из которых имеет специфическую структуру для распознавания антигена, форму которого заранее предположить невозможно?

В практических целях мы обычно полагаем, что млекопитающее может образовать от 10^6 до 10^8 различных типов антител. Каждое антитело — это иммуноглобулиновый тетрамер, состоящий из двух идентичных **легких цепей (L)** и двух идентичных **тяжелых цепей (Н)**. Если считать, что для создания антитела каждая легкая цепь может объединяться с каждой тяжелой цепью, то для 10^6–10^8 потенциальных антител необходимо 10^3–10^4 различных легких и столько же различных тяжелых цепей.

Существует два типа легких цепей и около 10 типов тяжелых. Разные классы иммуноглобулинов выполняют разные эффекторные функции. Класс иммуноглобулинов определяется константным участком тяжелой цепи, который и выполняет эффекторную функцию (см. рис. 23.17).

Структура тетрамера-иммуноглобулина показана на РИС. 23.4. Легкие и тяжелые цепи имеют в целом сходную организацию. При этом каждая из них состоит из двух принципиально различных районов: N-концевого **вариабельного участка (V-участок)** и С-концевого **константного участка (С-участок)**. Исходно их определили путем сравнения последовательностей аминокислотных остатков разных цепей иммуноглобулинов. Как предполагают названия, вариабельный участок имеет большое разнообразие последовательностей у разных белков-иммуноглобулинов, в то время как константный участок обладает более выраженной консервативностью.

Для образования отдельных доменов белка-иммуноглобулина соответствующие районы легких и тяжелых цепей соединяются.

Вариабельный домен (V) образуется путем ассоциации вариабельных участков легкой и тяжелой цепи. *V-домен отвечает за распознавание антигена.* Иммуноглобулин имеет Y-подобную структуру, в которой плечи этой буквы Y идентичны и каждое несет одну копию V-домена. Образование V-доменов с различными специфичностями дает возможность отвечать на разнообразные антигены. В целом число вариабельных участков для белков как легких, так и тяжелых цепей измеряется сотнями. *Таким образом, эти белки максимально разнообразны в участках, ответственных за связывание антигена.*

Число константных участков значительно меньше, чем число вариабельных, — обычно от одного до десяти С-участков для каждого типа цепи. Константные участки в иммуноглобулиновом тетрамере ассоциированы друг с другом таким образом, чтобы создавать отдельные С-домены. Первый домен — результат ассоциации единственного константного участка

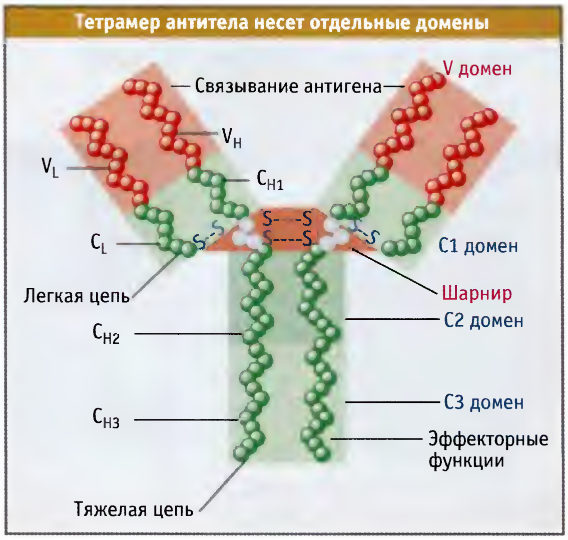


РИС. 23.4 **Тяжелая и легкая цепи объединяются для образования иммуноглобулина, имеющего несколько отдельных доменов**

легкой цепи (C_L) с C_{H1}-частью константного участка тяжелой цепи. Две копии этого домена создают плечи Y-подобной молекулы. Путем объединения С-районов тяжелых цепей образуются дополнительные С-домены, причем число этих доменов в одной молекуле может варьировать в зависимости от типа тяжелой цепи.

Сравнивая характеристики вариабельных и константных участков, мы видим главную дилемму в структуре гена иммуноглобулина. Каким образом в геноме закодирован целый набор белков, в каждом из которых любая отдельная полипептидная цепь должна иметь один из менее 10 возможных С-участков, но при этом также иметь один из нескольких сотен возможных V-участков? Оказывается, что число кодирующих последовательностей для каждого участка отражает его вариабельность. Существует множество генов, кодирующих V-участки, но лишь несколько генов для С-участков.

В этом контексте *понятие «ген» означает последовательность ДНК, кодирующую отдельную часть иммуноглобулинового полипептида* (часть тяжелой или легкой цепи). Таким образом, V-гены кодируют вариабельные участки, а С-гены кодируют константные участки, хотя *никакой из этих генов не экспрессируется как независимая единица.* V-ген должен физически соединиться с С-геном, чтобы получилась последовательность, способная экспрессироваться в виде аутентичной легкой или тяжелой цепи. В этой системе два «гена» кодируют один полипептид. Во избежание недопонимания мы будем называть эти последовательности «генными сегментами», а не «генами».

Последовательности, кодирующие легкие и тяжелые цепи, соединяются таким же образом: *любой из множества V-генных сегментов может соединиться с любым из нескольких С-генных сегментов.*

Эта **соматическая рекомбинация** происходит *в В-лимфоците, нарабатывающем антитело.* Основная роль в образовании огромного разнообразия иммуноглобулинов принадлежит большому числу доступных V-генных сегментов. Однако многообразие не полностью закодировано в геноме; определенные вариации возникают в процессе образования функционального гена.

В целом все это описание можно применить и к образованию функциональных генов, кодирующих белковые цепи Т-клеточных рецепторов. На поверхности Т-клеток содержатся два типа рецепторов — один, состоящий из двух типов цепей, называемых α и β, и второй, состоящий из γ и δ цепей. Подобно генам иммуноглобулинов, гены, кодирующие цепи Т-клеточных рецепторов, состоят из отдельных частей, включающих V- и С-области, объединяющиеся в активной Т-клетке (см. разд. 23.18 «Т-клеточные рецепторы схожи с иммуноглобулинами»).

Таким образом, основной момент синтеза иммуноглобулинов состоит в том, что *объединение V-генных и С-генных сегментов в клетках, производящих иммуноглобулины (или Т-клеточные рецепторы), отличает эти клетки от всех остальных соматических клеток или клеток зародышевой линии.*

Создание гена функционального иммуноглобулина или гена Т-клеточного рецептора можно рассматривать как процесс, описанный Ламарком: изменение генома в ответ на определенную особенность фенотипа (антиген). При рождении организм не имеет функционального гена для производства конкретного антитела или Т-клеточного рецептора. Он имеет большое число V-генных сегментов и гораздо меньшее число С-генных сегментов. Впоследствии активный ген создается из этих частей для производства антитела, способного реагировать с антигеном. Теория клональной селекции предполагает, что эта реорганизация ДНК осуществляется *до встречи с антигеном*, а затем происходит *селекция* клеток, несущих белок, способный связываться с антигеном. Весь процесс происходит в соматических клетках и не затрагивает клетки зародышевой линии; таким образом, ответ на антиген не передается потомству организма.

Существует два семейства легких цепей иммуноглобулинов, κ и λ, и одно семейство, содержащее все типы тяжелых цепей (H). Каждое семейство расположено на разных хромосомах и состоит только из своего набора V-генных и С-генных сегментов. Это называют *зародышевым паттерном*, и такая структура находится в клетках зародышевой линии и в соматических клетках всех других линий, за исключением клеток иммунной системы.

В клетке, экспрессирующей антитело (каждая из его цепей — одна легкая цепь (κ или λ) и одна тяжелая цепь) кодируется одним интактным геном. Путем рекомбинации, соединяющей V-генный сегмент и С-генный сегмент, создается активный ген, состоящий из экзонов, точно соответствующих функциональным доменам

белка. Интроны удаляются обычным способом в ходе сплайсинга РНК.

Рекомбинация между V- и C-генными сегментами для образования функциональных локусов происходит в популяции незрелых лимфоцитов. У B-лимфоцита обычно бывает лишь одна продуктивная реорганизация генных сегментов легких цепей (κ или λ) и одна — генных сегментов тяжелых цепей. Сходным образом, и T-лимфоцит продуктивно реорганизует α-ген и β-ген или один δ-ген и один γ-ген. Структура антитела или T-клеточного рецептора определяется конкретной конфигурацией соединившихся V- и C-генных сегментов в клетке, их производящей.

Принципы, благодаря которым образуются функциональные гены, одни и те же в каждом семействе, но есть различия в деталях организации V-генных и C-генных сегментов и, соответственно, в реакциях рекомбинации между ними. Кроме V- и C-генных сегментов в функциональные соматические локусы включаются и другие короткие последовательности ДНК (в том числе J-сегменты и D-сегменты).

23.4 Легкие цепи образуются путем единственного акта рекомбинации

Ключевые понятия:

- Легкая цепь лямбда формируется с помощью единственного акта рекомбинации между V-геном и сегментом J-C.
- V-генный сегмент имеет лидерный экзон, интрон и последовательность, кодирующую вариабельную область.
- J-C генный сегмент состоит из короткого J-кодирующего экзона, интрона и C-кодирующей области.
- Легкая цепь каппа формируется путем единственного акта рекомбинации между V-генным сегментом и одним из пяти J-сегментов, предшествующих C-гену

Легкая цепь λ собирается из двух частей, как показано на рис. 23.5. V-генный сегмент состоит из лидерного экзона (L), отделенного от вариабельного сегмента (V) единственным интроном. C-генный сегмент состоит из J-сегмента, отделенного интроном от константного (C) экзона.

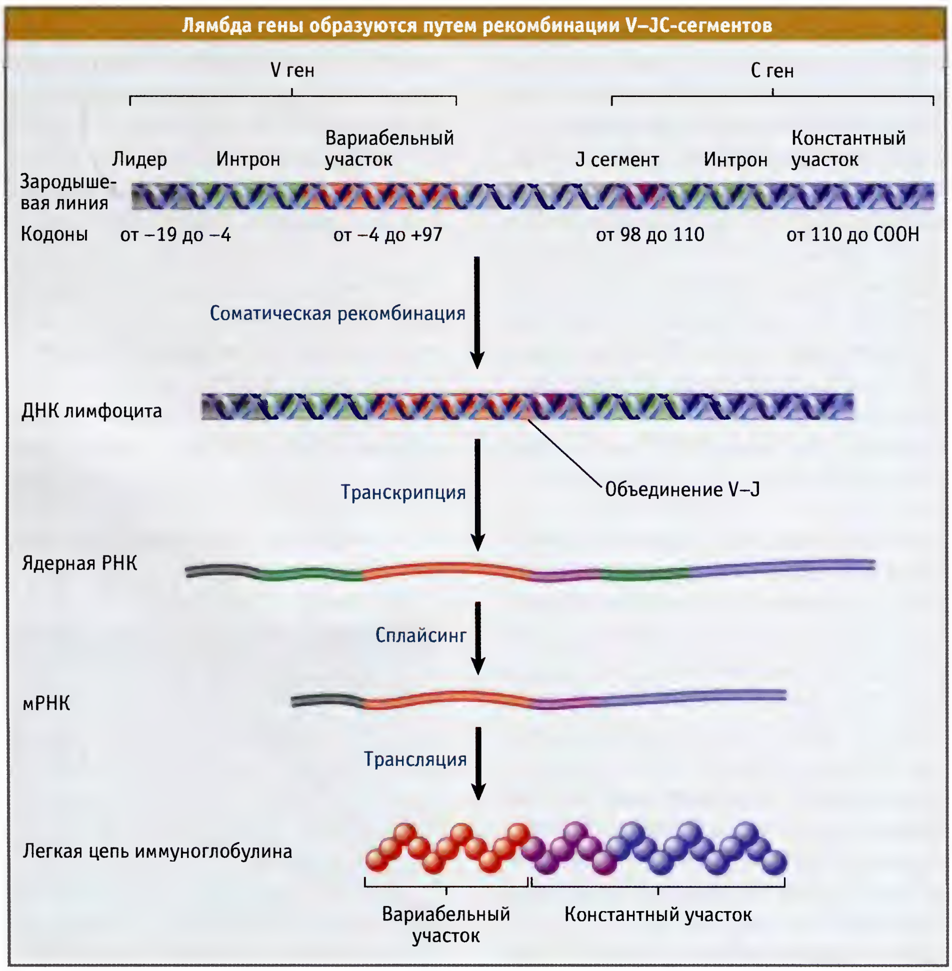


РИС. 23.5 C-генному сегменту лямбда предшествует J-сегмент; таким образом, V-J рекомбинация создает функциональный ген легкой цепи лямбда

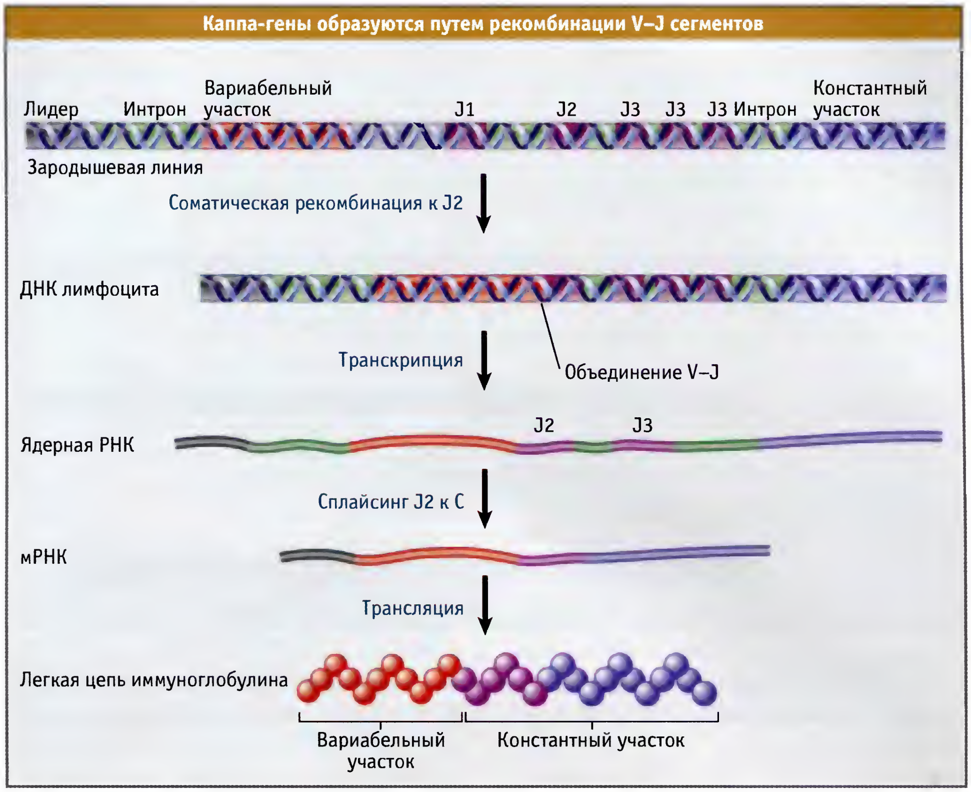


РИС. 23.6 С-генному сегменту каппа предшествует множество J-сегментов в клетках зародышевого пути. V-J соединение может быть между любым из этих J-сегментов, после чего происходит сплайсинг к С-сегменту во время РНК-процессинга

J-сегмента (от *англ.* joining — соединение) обозначает район, с которым соединяется V-сегмент. Таким образом, реакция объединения не происходит непосредственно между V- и С- сегментами, а происходит через J-сегмент; и когда мы обсуждаем соединение V- и С-сегментов в случае легких цепей, мы в действительности имеем в виду соединение V-JC.

J-сегмент довольно короткий и кодирует последние несколько аминокислотных остатков вариабельного района (как определили путем аминокислотного секвенирования). В интактном гене, образованном рекомбинацией, V-J сегмент представляет собой один экзон, кодирующий весь вариабельный участок.

Последовательность реакций объединения сегментов цепи κ показана на РИС. 23.6. Легкая цепь κ собирается также из двух частей, но в этом случае есть некоторые отличия в организации С-генного сегмента. Группа из пяти J-сегментов распределена на участке от 500 до 700 пн и отделена от экзона Cκ интроном от 2 до 3 тпн. У мыши центральный J-сегмент не функционален (ψJ3). Vκ-сегмент может соединиться с любым из J-сегментов.

Какой бы J-сегмент ни использовался, он становится конечной частью интактного вариабельного экзона. Все остальные J-сегменты слева от рекомбинирующего J-сегмента удаляются (J1 на рисунке был удален). Все остальные J-сегменты справа от рекомбинирующего J-сегмента воспринимаются как часть интрона между вариабельным и константным экзонами (J3 включается в интрон, который на рисунке вырезается).

Все функциональные J-сегменты несут сигнал на левой границе, дающий им возможность рекомбинировать с V-сегментом; также они имеют сигнал на правой границе, который может использоваться для сплайсинга к С-экзону. Какой бы J-сегмент ни соединялся с V-сегментом, после соединения используется его сигнал сплайсинга для процессинга РНК.

23.5 Тяжелые цепи образуются путем двух актов рекомбинации

Ключевые понятия:

- Для формирования тяжелых цепей в рекомбинации используются такие элементы, как V-ген, D-сегмент и J-С-генный сегмент.
- Первый раунд рекомбинации объединяет D- и J-С-сегменты.
- Второй раунд рекомбинации соединяет V- с D-J-С-сегментами
- С-сегмент состоит из нескольких экзонов.

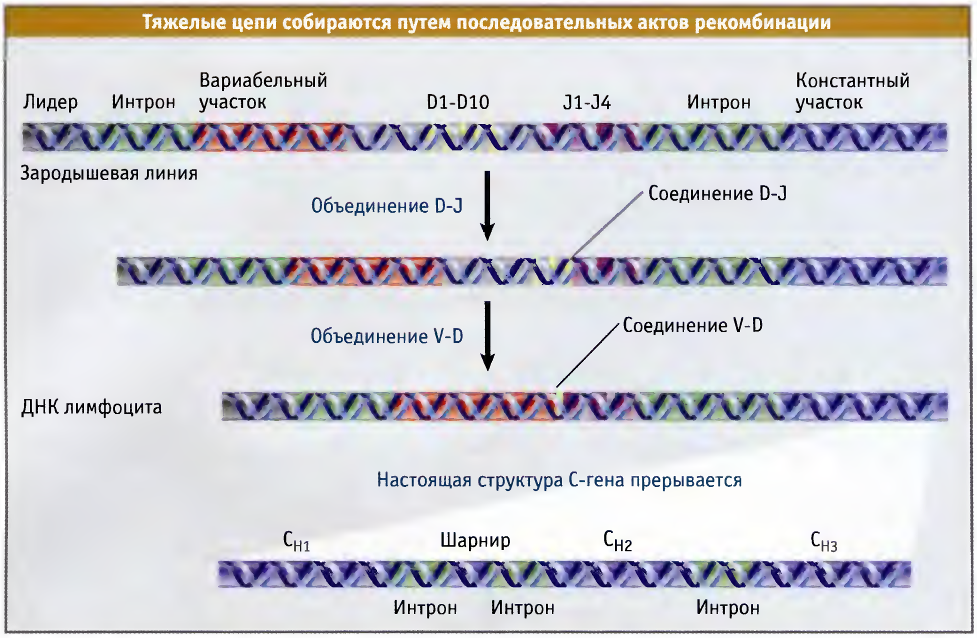


РИС. 23.7 Гены тяжелых цепей образуются путем последовательных реакций соединения. Сначала D-сегмент соединяется с J-сегментом, а затем V-генный сегмент соединяется с D-сегментом

В образовании тяжелой цепи участвует дополнительный сегмент. **D-сегмент** (от *англ.* diversity — разнообразие) был открыт в виде дополнительных от двух до тринадцати аминокислотных остатков в белке между последовательностями, кодируемыми V-сегментом и J-сегментом. Участок с более чем 10 D-сегментами находится на хромосоме между множеством V_H-сегментов и четырьмя J_H-сегментами.

Объединение V-D-J происходит в две стадии, как показано на РИС. 23.7. Сначала один из D-сегментов рекомбинирует с J_H-сегментом; а V_H-сегмент затем рекомбинирует с DJ_H объединенным сегментом. Это соединение ведет к экспрессии смежного C_H-сегмента (состоящего из нескольких экзонов). (Мы обсудим использование различных C_H-генных сегментов в разд. 23.12 «Рекомбинация ДНК вызывает переключение классов иммуноглобулинов»; сейчас мы просто описываем реакцию в терминах соединения одного из нескольких J-сегментов, предшествующих C_H-генному сегменту.)

D-сегменты организованы в виде тандемной последовательности. Мышиный локус тяжелых цепей содержит двенадцать D-сегментов различной длины; этот локус у человека несет около 30 D-сегментов (но не все из них активны). Пока не выяснен механизм, обеспечивающий участие *того же самого* D-сегмента в D-J и V-D реакциях соединения. (При обсуждении объединения V- и C-сегментов тяжелых цепей мы допускаем, что этот процесс осуществляется путем V-D и D-J реакций соединения.)

V-генные сегменты всех трех семейств иммуноглобулинов схожи в организации. Первый экзон кодирует сигнальную последовательность (необходимую для присоединения к мембране), а второй экзон кодирует основную часть собственно вариабельного участка (длиной <100 кодонов). Оставшуюся часть кодирует D-сегмент (только в H-семействе) и J-сегмент (во всех трех семействах).

Структура константного участка зависит от типа цепи. У легких цепей κ и λ константный участок кодируется единственным экзоном (который становится третьим экзоном образованного активного гена). У H-цепей константный участок кодируется несколькими экзонами; в соответствии со структурой белковой цепи на рис. 23.4 разные экзоны кодируют участки: C_{H1}, шарнир, C_{H2}, C_{H3}. Каждый CH-экзон составляет около 100 кодонов; шарнирный участок короче. Интроны обычно относительно малы (~300 пн).

23.6 Рекомбинация создает огромное разнообразие антител

Ключевые понятия:

- Локус легких цепей может произвести более 1000 различных цепей путем комбинации 300 V-генов с четырьмя-пятью C-генами.
- H-локус может произвести более 4000 цепей, комбинируя 300 V-генов, 20 D-сегментов и четыре J-сегмента.

Теперь нам нужно рассмотреть различные типы V- и C-генных сегментов, чтобы увидеть, какое многообразие может быть получено с учетом наличия серии кодирующих последовательностей в клетках зародышевого пути. В каждом генном семействе легких цепей иммуноглобулинов множество V-генных сегментов сцеплено с гораздо меньшим числом C-генных сегментов.

На РИС. 23.8 показано, что локус λ несет около 6 C-генных сегментов, каждому из которых предшествует свой собственный J-сегмент. Локус λ у мыши содержит значительно меньше вариантов генных сегментов, чем у человека. Основное отличие в том, что у мыши есть только два $V_λ$-генных сегмента; каждый сцеплен с двумя J-C-сегментами. Из четырех $C_λ$-генных сегментов один неактивен. В прошлом геном мыши подвергся катастрофической делеции большинства своих $V_λ$-генных сегментов зародышевого пути.

На РИС. 23.9 показано, что локус κ несет только один C-генный сегмент, но ему предшествуют пять J-сегментов (один из которых неактивен). $V_κ$-генные сегменты образуют большой кластер на хромосоме выше сегмента константного района. У человека этот кластер имеет два участка. Непосредственно предшествующий Cκ-генному сегменту участок в 600 пн содержит пять Jκ-сегментов и 40 Vκ-генных сегментов. От еще одной группы в 36 Vκ-сегментов этот участок отделен промежутком в 800 пн.

Vκ-генные сегменты можно разделить на семейства, члены в которых будут иметь более 80% сходства аминокислотных последовательностей. Мышиное разнообразие особенно обширно (~1000 генов): существует ~18 семейств Vκ, варьирующих в размере от 2 до 100 членов. Подобно другим семействам родственных генов, родственные V-генные сегменты формируют подгруппы, возникшие в результате дупликации и дивергенции древних членов-предшественников. При этом многие из V-сегментов представляют собой неактивные псевдогены, и, по-видимому, <50 используются для образования иммуноглобулинов.

Каждый отдельный лимфоцит образует *либо* κ-, *либо* λ-цепь для ассоциации с тяжелой цепью. У человека ~60% образующихся легких цепей принадлежат к семейству κ, и ~40% — к семейству λ. У мыши 95% B-клеток экспрессируют легкую цепь κ, вероятно, из-за небольшого числа λ-генных сегментов.

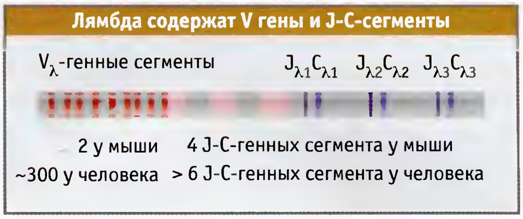


РИС. 23.8 Семейство лямбда состоит из V-генных сегментов, сцепленных с небольшим числом J-C-генных сегментов

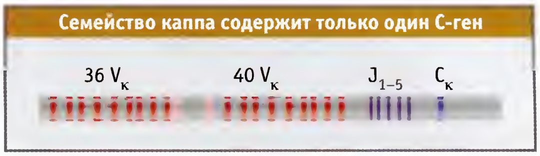


РИС. 23.9 Семейство каппа у мыши и у человека состоит из V-генных сегментов, сцепленных с пятью J-сегментами, соединенных с единственным C-сегментом

Единственный локус тяжелых цепей у человека состоит из нескольких отдельных участков, как показано на РИС. 23.10. У мыши схожая ситуация, хотя у них больше V_H-генных сегментов, меньше D- и J-сегментов, и есть небольшие отличия в числе и организации C-генных сегментов. Первый 3′-участник V_H-кластера отделен всего лишь 20 тпн от первого D-сегмента. D-сегменты распределены на ~50 тпн и сопровождаются кластером J-сегментов. В следующих 220 тпн содержатся все C_H-генные сегменты. Существует девять функциональных C_H-генных сегментов и два псевдогена. Такая организация предполагает, что γ-генный сегмент в прошлом был дуплицирован, дав участок γ–γ–ε–α, после чего вся эта группа также была дуплицирована.

Насколько многообразие V-участка у белков-иммуноглобулинов обусловлено именно разнообразием информации в клетках зародышевой линии? Путем комбинирования любого из ~50 V-генных сегментов с любым из четырех-пяти J-сегментов типичный локус легких цепей может дать около 250 различных цепей.

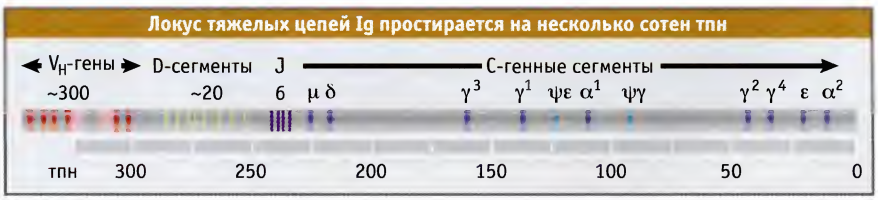


РИС. 23.10 У человека единственный генный кластер содержит всю информацию, необходимую для формирования тяжелых цепей

В локусе H-цепей заключается еще большее многообразие; путем комбинации каждого из ~50 V_H-генных сегментов, 20 D-генных сегментов и четырех J-сегментов потенциально геном может образовать 4000 вариабельных участков для объединения с C_H-генным сегментом. У млекопитающих это исходная точка для формирования многообразия, но существуют дополнительные механизмы, вносящие дальнейшие изменения. *Исследуя близкородственные варианты иммуноглобулинов, часто можно найти больше белков, чем существующее число соответствующих V-генных сегментов.* Новые члены создаются путем соматических изменений в индивидуальных генах во время или после процесса рекомбинации (см. разд. 23.14 «Соматические мутации создают еще большее многообразие антител у мышей и человека»).

23.7 При рекомбинации генов иммунной системы используются два типа консенсусных последовательностей

Ключевые понятия:

- Консенсусная последовательность, используемая для рекомбинации, — это гептамер, отделенный интервалом в 12 пн или 23 пн от нонамера.
- Рекомбинация происходит между двумя консенсусными последовательностями, которые имеют разные интервалы.

При сборке генов легких и тяжелых цепей используются одни и те же механизмы (хотя количество составных частей отличается). На границах всех сегментов, участвующих в реакциях соединения, найдены одни и те же консенсусные последовательности. Каждая консенсусная последовательность состоит из гептамера, отделенного 12 или 23 пн от нонамера.

На рис. 23.11 показаны взаимодействия между консенсусными последовательностями в мышиных локусах генов иммуноглобулинов. В κ-локусе каждый Vκ-сегмент сопровождается консенсусной последовательностью с 12-нуклеотидным интервалом (спейсером). Каждому Jκ-сегменту предшествует консенсусная последовательность с 23-нуклеотидным спейсером. V и J консенсусные последовательности имеют инвертированную ориентацию. В λ-локусе каждый Vλ-сегмент сопровождается консенсусной последовательностью с 23-нуклеотидным спейсером; а каждому Jλ-сегменту предшествует консенсусная последовательность с 12-нуклеотидным спейсером.

Правило, по которому происходят реакции соединения, гласит: *консенсусная последовательность с одним типом спейсера может соединиться с консенсусной последовательностью, несущей только другой тип спейсера.* Консенсусные последовательности у V- и J-сегментов могут располагаться в любом порядке; таким образом, разные спейсеры не несут никакой направляющей информации, а служат лишь для предотвращения рекомбинации V- или J-сегмента с таким же сегментом.

Такая модель подтверждается и структурой компонентов генных сегментов тяжелых цепей. Каждый генный сегмент V_H сопровождается консенсусной последовательностью с 23-нуклеотидным типом спейсера. D-сегмент фланкирован со всех сторон консенсусной последовательностью с 12-нуклеотидным спейсером. J_H-сегменту предшествует консенсусная последовательность с 23-нуклеотидным спейсером. Итак, V-генный сегмент должен соединиться с D-сегментом, а D-сегмент должен быть соединен с J-сегментом. V-генный сегмент не может быть соединен напрямую с J-сегментом, поскольку оба они несут один и тот же тип консенсусной последовательности.

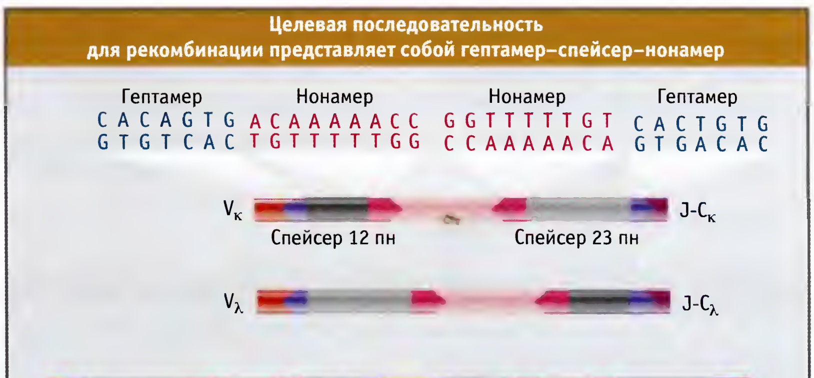


Рис. 23.11 У каждой пары рекомбинирующих сайтов консенсусная последовательность присутствует в инвертированной ориентации. Один член из каждой пары имеет между компонентами интервал в 12 пн; второй — в 23 пн

Спейсеры между компонентами консенсусных последовательностей почти соответствуют одному или двум виткам двойной спирали. Это может отражать их пространственное расположение во время реакции рекомбинации. Например, белок (или белки) рекомбинации может подходить к ДНК с одной стороны, подобно тому, как РНК-полимераза и репрессоры приближается к распознаваемым элементам, таким как промоторы и операторы.

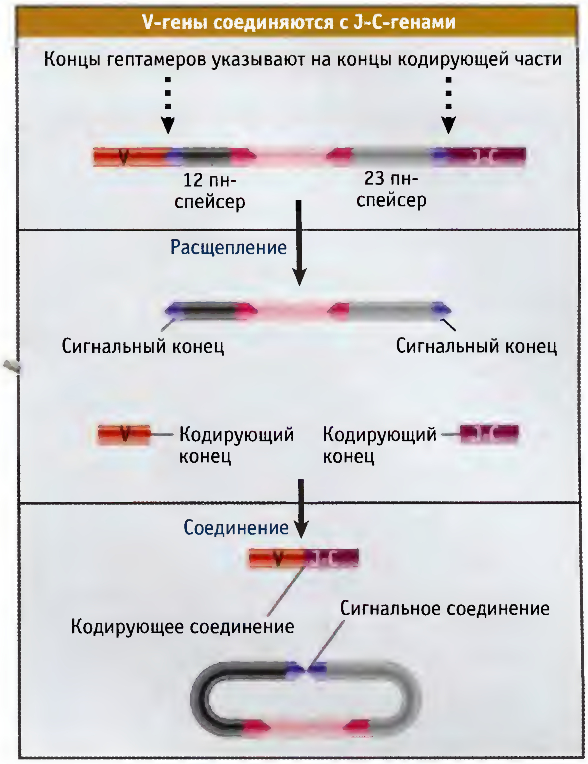


РИС. 23.12 Разрыв и воссоединение в консенсусных последовательностях приводят к образованию генов иммунглобулинов

23.8 Рекомбинация вызывает делеции или инверсии

Ключевые понятия:

- Рекомбинация происходит путем двухцепочечных разрывов в районе гептамеров двух консенсусных последовательностей.
- Сигнальные концы фрагмента между точками разрывов обычно соединяются друг с другом с образованием кольцевого вырезаемого фрагмента.
- Кодирующие последовательности ковалентно соединяются друг с другом: V с J-C (L-цепь), а D с J-C и V- с D-J-C (H-цепь).
- Если рекомбинирующие гены инвертированы друг относительно друга, то вместо делеции кольцевого фрагмента происходит инверсия.

Рекомбинация компонентов генов иммуноглобулинов завершается физической реорганизацией последовательностей, включающей в себя разрыв и воссоединение цепей, однако механизм отличается от гомологичной рекомбинации. Основной принцип реакции для легкой цепи λ показан на РИС. 23.12. (Реакция в локусе тяжелых цепей аналогична, только там происходят два события рекомбинации: сначала D-J, затем V-DJ.)

Разрыв и воссоединение происходят как отдельные реакции. Двухцепочечный разрыв производится в районе гептамеров на конце кодирующего участка. Это высвобождает целый фрагмент между V-генным сегментом и J-C-генным сегментом; отрезанные концы этого фрагмента называют **сигнальными концами**. Свободные концы V- и J-C-локусов называют **кодирующими концами**. Два кодирующих конца ковалентно соединяются; это соединение связывает V- и J-сегменты. Если бы два сигнальных конца также соединялись, то вырезаемый фрагмент образовывал бы кольцевую молекулу.

Мы показали, что V- и J-C-локусы имеют одинаковую ориентацию. В результате, разрезание в каждой консенсусной последовательности высвобождает участок между ними в виде линейного фрагмента. Если сигнальные концы соединяются, то он превращается в кольцевую молекулу, как показано на рис. 23.12. Делеция с вырезанием кольцевой молекулы — это преобладающий способ рекомбинации в локусах иммунглобулинов и TCR.

В некоторых исключительных случаях V-генный сегмент имеет обратную ориентацию на хромосоме относительно J-C-локуса. В таком случае разрыв и воссоединение инвертирует промежуточный материал (вместо его делеции). Результат инверсии вместо делеции такой же, как было показано ранее для гомологичной рекомбинации между прямыми и инвертированными повторами на рис. 21.9 и рис. 21.10, с одной оговоркой: при рекомбинации с инвертированным V-генным сегментом сигнальные концы *должны быть* соединены, поскольку иначе в локусе останется разрыв. Инверсии встречаются при рекомбинации TCR, а также (иногда) в локусе легких цепей κ.

23.9 Аллельное исключение инициируется продуктивной реорганизацией генома

Ключевые понятия:

- Рекомбинация, ведущая к образованию интактного гена иммуноглобулина, продуктивна, если образующийся ген способен экспрессировать активный белок.
- Продуктивная рекомбинация генетического материала предотвращает дальнейшую реорганизацию, в то время как непродуктивная продолжает перестройку.
- Аллельное исключение применяется отдельно к легким цепям (только одна каппа или лямбда цепь может быть продуктивно реорганизована) и к тяжелым цепям (только одна тяжелая цепь подвергается продуктивной реорганизации).

Каждая В-клетка экспрессирует один тип легкой цепи и один тип тяжелой цепи, поскольку в конкретном лимфоците происходит только одна продуктивная реорганизация каждого типа для образования единственного гена легкой цепи и единственного гена тяжелой цепи. В каждом акте участвуют гены только одной гомологичной хромосомы, и в результате *аллели с другой хромосомы не экспрессируются в той же клетке*. Это явление называют **аллельным исключением**.

Аллельное исключение затрудняет анализ соматической рекомбинации. Зонд, комплементарный участку, реорганизованному на одной гомологичной хромосоме, также будет детектировать аллельные последовательности на другой хромосоме. Таким образом, нам придется анализировать разную судьбу обеих хромосом.

Обычный паттерн, образованный реорганизованным активным геном, можно интерпретировать в терминах делеции материала между рекомбинирующими V- и С-локусами.

В активных клетках можно встретить два типа генной организации:

- Зонды на активный ген могут показать наличие одной реорганизованной копии и одной копии линии зародышевого пути. Следовательно, мы предполагаем, что объединение произошло на одной хромосоме, в то время как другая хромосома осталась неизмененной.
- Могут быть обнаружены два различных паттерна реорганизации, указывая на то, что обе хромосомы подверглись независимым актам реорганизации. В некоторых из этих случаев материал между V- и С-генными сегментами полностью отсутствует в клеточной линии. Такое положение дел легче всего объяснить независимыми делециями (в результате рекомбинации) на каждой хромосоме.

В случае, когда на обеих хромосомах не остается паттерна линии зародышевого пути, обычно только одна из них оказывается прошедшей **продуктивную реорганизацию** с образованием функционального гена. Вторая хромосома подверглась **непродуктивной реорганизации**; такая реорганизация может иметь несколько форм, однако в каждом случае последовательность гена не может быть экспрессирована в виде белка-иммуноглобулина. (Последовательность может быть неполной, например, если D-J соединение произошло, а V-D соединения не было; либо последовательность может быть нарушена, когда процесс соединения был завершен, но не оказался успешным в образовании гена, кодирующего функциональный белок.)

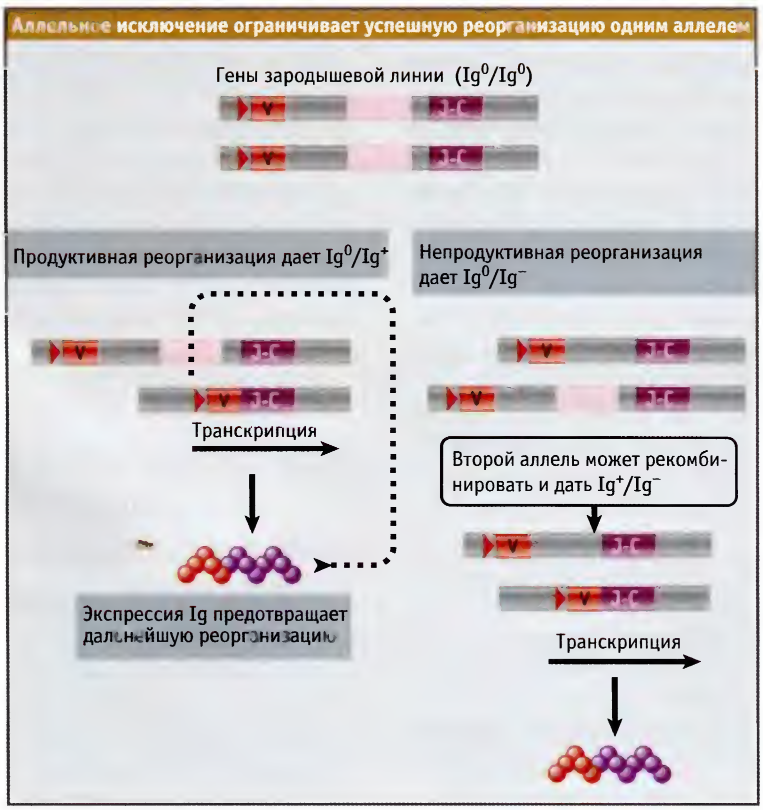


РИС. 23.13 Успешная реорганизация с образованием активной легкой или тяжелой цепи подавляет дальнейшую реорганизацию того же типа и приводит к аллельному исключению

Одновременное наличие продуктивной и непродуктивной реорганизации предполагает наличие обратной связи для контроля процесса рекомбинации. Модель обратной связи представлена на РИС. 23.13. Предположим, что каждая клетка вступает в процесс с двумя локусами в нереорганизованной конфигурации линии зародышевого пути Ig^0. Каждый из этих локусов может быть реорганизован с образованием продуктивного гена Ig^+ или непродуктивного гена Ig^-.

Если реорганизация продуктивна, синтез активной цепи предотвращает реорганизацию другого аллеля. Тогда активная клетка имеет конфигурацию Ig^0/Ig^+.

Если реорганизация оказалась непродуктивной, образуется клетка с конфигурацией Ig^0/Ig^-. При этом нет препятствий для реорганизации оставшегося аллеля с конфигурацией зародышевого пути. Если эта реорганизация оказывается продуктивной, экспрессирующая клетка имеет конфигурацию Ig^+/Ig^-. И снова присутствие активной цепи подавляет дальнейшую возможную реорганизацию.

После двух последовательных непродуктивных реорганизаций образуется клетка Ig^-/Ig^-. Иногда клетка Ig^-/Ig^- может снова подвергнуться реорганизации. Наблюдаемый в некоторых случаях паттерн ДНК мог возникнуть только путем последовательных реорганизаций.

Суть модели в том, что клетка пытается рекомбинировать V- и C-генные сегменты до тех пор, пока не будет достигнута продуктивная реорганизация. Как только образуется активная цепь, происходит аллельное исключение путем супрессии последующих актов реорганизации. Существование такого механизма *in vivo* показано с помощью созданных трансгенных мышей, у которых клетки зародышевого пути несут уже реорганизованный ген иммуноглобулина. Экспрессия трансгена в B-клетках подавляет реорганизацию эндогенных генов.

Аллельное исключение независимо для локусов тяжелых и легких цепей. Гены тяжелых цепей обычно реорганизуются первыми. Аллельное исключение для легких цепей должно применяться в равной мере к обоим семействам (поскольку клетки могут иметь либо κ, либо λ активную легкую цепь). Вероятно, клетки сначала реорганизуют гены семейства κ, а гены семейства λ реорганизуют, только если обе попытки с генами κ оказались неуспешными.

В этих сериях событий есть интересный парадокс. В реакциях рекомбинации участвуют одни и те же консенсусные последовательности и одна и та же V(D)J-рекомбиназа в H-, κ- и λ-локусах, однако эти три локуса реорганизуются в определенной последовательности. Что обеспечивает реорганизацию тяжелой цепи до реорганизации легких цепей, а реорганизацию κ-цепей до реорганизации λ-цепей? Эти локусы могут быть доступны для ферментов в различное время, возможно, в результате транскрипции. Транскрипция осуществляется даже до реорганизации, хотя, разумеется, продукты ничего не кодируют. Транскрипция может изменять структуру хроматина, обеспечивая доступ фермента к консенсусным последовательностям для рекомбинации.

23.10 Белки RAG катализируют разрыв и воссоединение ДНК

Ключевые понятия:

- Для реакции расщепления необходимо и достаточно активности белков RAG.
- RAG1 распознает консенсусную последовательность-нонамер для рекомбинации. RAG2 связывается с RAG1 и расщепляет ДНК в районе гептамера.
- Реакция напоминает события, происходящие при разрешении цепей в результате работы топоизомеразы при транспозиции.
- На конце, несущем кодирующую последовательность, образуется промежуточный продукт в виде шпильки; раскрывание этой шпильки приводит к внедрению дополнительных оснований (P-нуклеотидов) в образующийся ген.
- Дезоксинуклеозид-трансфераза вводит добавочные N-нуклеотиды в кодирующий конец.
- Кодон в месте объединения V-(D)-J сегментов исключительно вариабелен и кодирует аминокислоту 96 в районе антиген-связывающего участка.
- Двухцепочечные разрывы в месте объединения кодирующих последовательностей репарируются той же системой, что действует при соединении негомологичных концов поврежденной ДНК.
- Энхансер в C-гене активирует промотор V-гена после рекомбинации (в случае образования интактного гена иммуноглобулина).

Белки RAG1 и RAG2 необходимы и достаточны для расщепления ДНК для V(D)J при рекомбинации. Их кодируют два гена, отделенные друг от друга <10 тпн на хромосоме. Трансфекция этих генов в фибробласты приводит к тому, что соответствующая ДНК подвергается реакции соединения V(D)J. Мыши, не имеющие генов *RAG1* и *RAG2*, неспособны рекомбинировать гены иммуноглобулинов и T-клеточных рецепторов и в результате имеют только незрелые B- и T-лимфоциты. Белки RAG производят каталитические реакции расщепления и воссоединения ДНК, а также обеспечивают структурную основу, в рамках которой и происходят эти реакции.

RAG1 распознает сигналы гептамеров/нонамеров с соответствующими 12/23-нуклеотидными спейсерами и рекрутирует RAG2 в реакционный комплекс. Нонамер является сайтом для первичного распознавания, а гептамер указывает место расщепления.

Реакции, происходящие во время рекомбинации, показаны на РИС. 23.14. Комплекс делает одноцепочечный разрыв в каждом из мест соединения. Этот разрыв имеет 3′-OH и 5′-P концы. Свободный 3′-OH конец затем атакует фосфатную связь в соответствующей позиции *на другой стороне дуплекса*. Таким образом на кодирующем конце формируется шпилька, в которой 3′-конец одной цепи ковалентно соединен с 5′-концом другой цепи; при этом остается двухцепочечный разрыв на сигнальном конце.

Этот разрыв представляет собой реакцию трансэтерификации, в которой запасается энергия химических связей. Такая реакция напоминает реакции топоизомеризации, катализируемые белками-резольвазами из бактериальных траспозонов (см. разд. 21.9 «В транспо-

зиции TnA участвуют транспозаза и резольваза»). Сходство между этими реакциями поддерживается также тем, что белок RAG1 сходен с бактериальными инвертазами (инвертирующими специфические сегменты ДНК путем похожих реакций рекомбинации). Фактически белки RAG могут вставлять участки донорной ДНК, свободные концы которой содержат соответствующие сигнальные последовательности (гептамер-12/23-нуклеотидный спейсер-нонамер), в независимую целевую ДНК в реакции транспозиции *in vitro*. Эти факты могут говорить о том, что соматическая рекомбинация генов иммунной системы эволюционировала из предковых форм транспозонов. Кроме того, описанный механизм предполагает, что белки RAG отвечают за хромосомные транслокации, в которых Ig- или TCR-локусы соединяются с другими локусами.

Шпильки возле кодирующих концов создают субстрат для последующей стадии реакции. Если одноцепочечный разрыв вносится близко к шпильке, реакция плавления возле этого конца создает одноцепочечный выступ. Синтез комплементарной цепи превращает кодирующий конец в дуплекс. Эта реакция объясняет введение **Р-нуклеотидов** в кодирующий конец; они представляют собой несколько дополнительных пар оснований, расположенных в обратной ориентации относительно исходного кодирующего конца.

Между кодирующими концами могут быть введены несколько дополнительных оснований, по-видимому, в случайном порядке. Их называют **N-нуклеотидами**. Включение этих нуклеотидов происходит благодаря активности фермента дезоксинуклеозид-трансферазы (известно, что этот фермент активен в лимфоцитах) в районе свободного 3'-кодирующего конца, образующегося во время процесса соединения.

Изменения последовательности, таким образом, появляются вследствие энзиматических механизмов, участвующих в разрыве и воссоединении ДНК. В рекомбинации тяжелых цепей пары оснований теряются или вставляются в местах соединений V_H-D, D-J или в обоих случаях. Делеции также происходят при соединении

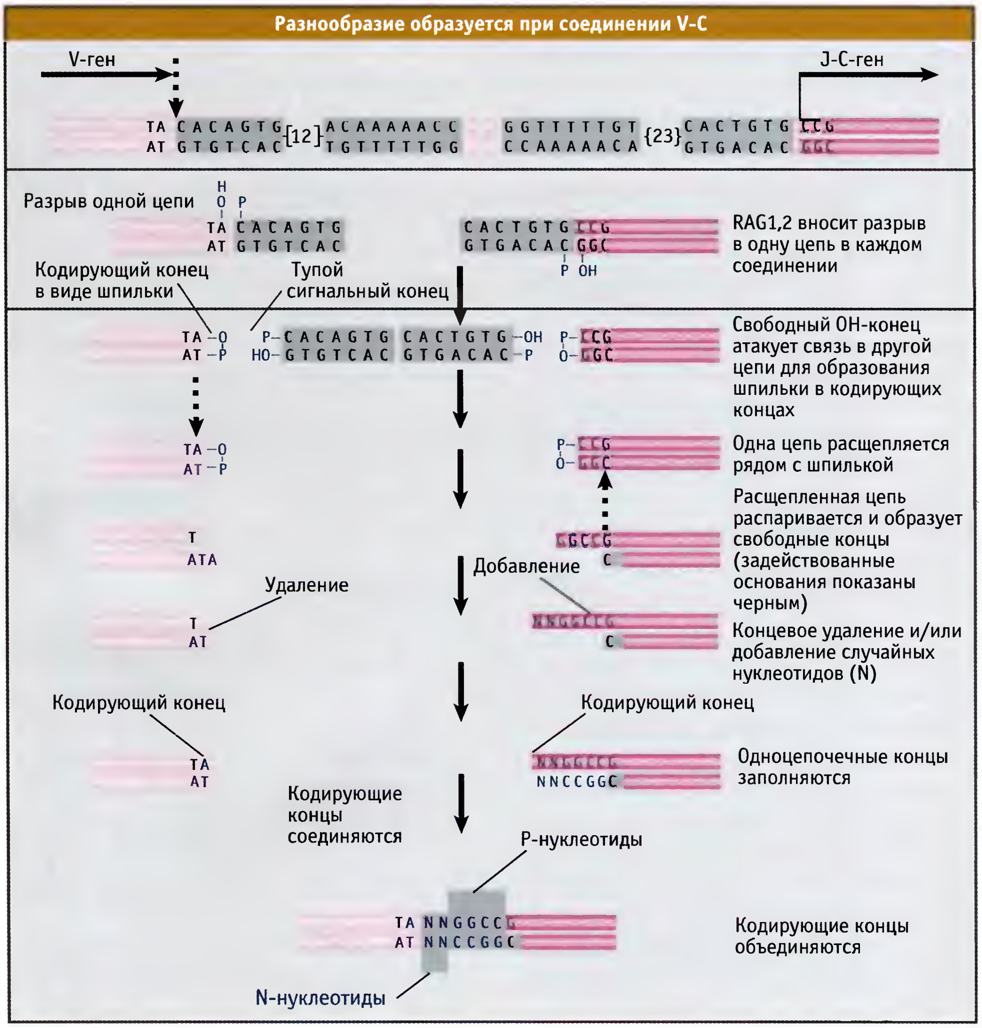


РИС. 23.14 При процессинге кодирующих концов вносятся изменения в местах соединений

Vλ-Jλ, а инсерции в этом месте встречаются редко. Изменения последовательностей оказывают влияние на последовательность аминокислотных остатков, закодированных в местах соединений V-D и D-J в тяжелых цепях или в месте соединения V-J в легких цепях.

Благодаря существованию нескольких различных механизмов обеспечивается ситуация, при которой место соединения может иметь последовательность нуклеотидов, отличающуюся от предсказанной (если исходить из непосредственного соединения кодирующих концов V-, D- и J-участков).

Изменения последовательности в местах соединений создают возможность для появления здесь большого разнообразия аминокислотных остатков. Интересно, что аминокислота в позиции 96 появляется вследствие реакции соединения V-J. Этот остаток образует часть антиген-связывающего сайта и также участвует в создании контакта между легкими и тяжелыми цепями. Таким образом, максимум разнообразия создается в местах контакта с целевым антигеном.

Изменение числа остатков нуклеотидов в местах соединений меняет рамку считывания. По-видимому, процесс соединения является случайным относительно рамки считывания; таким образом, вероятно, лишь одна треть собранных последовательностей сохраняет правильную рамку. Если V-J-участки соединяются таким образом, что J оказывается вне правильной рамки считывания, трансляция останавливается преждевременно из-за нонсенс-кодонов в этой области. Можно предположить, что образование нарушенных генов представляет собой ту цену, которую клетка вынуждена платить за увеличение разнообразия, приобретаемого при внесении изменений в последовательность в местах соединения участков.

Подобным образом — или даже в большей мере — создается разнообразие в реакциях соединения тяжелой цепи при участии D-сегмента. Относительно рамки считывания можно видеть такой же результат; в реакциях соединения образуются непродуктивные гены, в которых J- и C-сегменты выпадают из рамки предшествующего V-генного сегмента.

В реакциях соединения кодирующих концов используется тот же способ негомологичного соединения концов (NHEJ), с помощью которого в клетке репарируются двухцепочечные разрывы (см. разд. 20.11 «Системы репарации эукариотических клеток консервативны»). Начальные стадии реакции были изучены путем выделения промежуточных продуктов-интермедиатов из лимфоцитов мыши с мутацией *SCID* (от *англ.* severe combined immunodeficiency — тяжелый комбинированный иммунодефицит), приводящий к сильному падению уровня рекомбинации иммуноглобулиновых генов и генов TCR. Мыши, несущие *SCID*, накапливают аберрантные молекулы, оборванные в местах двухцепочечных разрывов у кодирующих концов, и, таким образом, дефектные в некоторых аспектах реакций соединения. Мутация *SCID* инактивирует ДНК-зависимую протеинкиназу (ДНК-ПК). Эта киназа направляется к ДНК белками Ku70 и Ku80, связывающимися с концами ДНК. ДНК-ПК фосфорилирует и, таким образом, активирует белок Artemis, который вносит разрыв по концам шпильки (также он имеет экзонуклеазную и эндонуклеазную активности, проявляющиеся в NHEJ пути). Лигирование выполняется ДНК-лигазой IV (при участии белка XRCC4). У людей мутации в белках Ku, XRCC4 или ДНК-лигазе IV найдены у пациентов с болезнями, вызываемыми дефектами репарации ДНК, что приводит к повышению чувствительности к облучению.

Какова же связь между соединением V- и C-генных сегментов и их активацией? Нереорганизованный V-генный сегмент неактивно представлен среди пула РНК. Однако после продуктивного присоединения V-генного сегмента к $C_κ$, полученная единица транскрибируется. Последовательность выше V-генного сегмента не изменяется в процессе объединения, и в результате *промотор должен быть одинаков в нереорганизованных, непродуктивно реорганизованных и продуктивно реорганизованных генах.*

Промотор лежит выше каждого V-генного сегмента, но он неактивен. Промотор активируется при перестановке его к C-участку. Этот эффект должен зависеть от вышележащих последовательностей. Какую же роль они могут играть? Энхансер внутри или выше C-генного сегмента активирует промотор возле V-генного сегмента. Этот энхансер тканеспецифичен; он активен только в B-клетках. Его существование позволяет предложить модель, показанную на рис. 23.15, согласно которой промотор V-генного сегмента активируется, когда его доставляют в область действия энхансера.

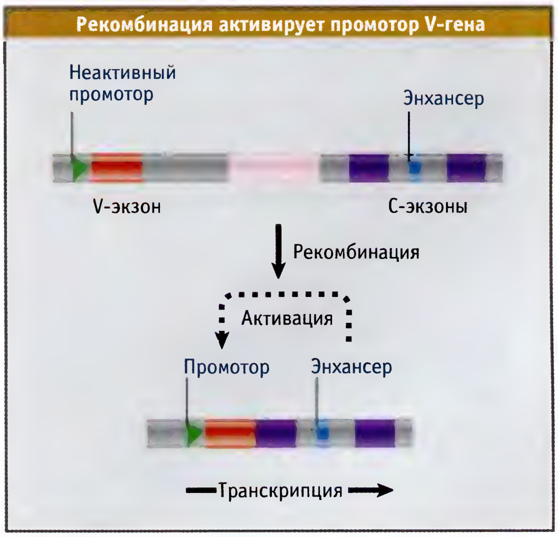


РИС. 23.15 Промотор V-генного сегмента неактивен до того момента, пока с помощью рекомбинации он не будет поставлен недалеко от энхансера в C-генном сегменте. Этот энхансер активен только в B-лимфоцитах

23.11 Процессинг РНК может изменить раннюю экспрессию тяжелой цепи

Ключевые понятия:

- Все лимфоциты начинают экспрессировать иммуноглобулины с мембраносвязанной формы IgM.
- Когда В-клетки дифференцируются, мембраносвязанная форма сменяется секреторной за счет сплайсинга РНК.

Период синтеза IgM, с которого начинается развитие лимфоцитов, можно разделить на две части, во время которых синтезируются различные версии μ-константного участка:

- В то время как стволовые клетки дифференцируются в пре-В-лимфоцит, синтезируется соответствующая легкая цепь, и IgM-молекула ($L_2\mu_2$) появляется на поверхности клетки. Эта форма IgM содержит версию μ_m константного участка (*m* показывает, что IgM находится на мембране). Мембранная локализация может отражать необходимость инициировать пролиферацию клеток в ответ на первичное распознавание антигена.
- Когда В-лимфоцит дифференцируется дальше (в плазматическую клетку), экспрессируется μ_s версия константного участка. IgM секретируется в виде пентамера IgM_5J, в котором J – это соединительный полипептид (не имеет отношения к J-участку), образующий дисульфидные связи с μ-цепями. Секреция белка сопровождается гуморальным ответом, как показано на рис. 23.1.

Версии тяжелой цепи μ_m и μ_s различаются только на С-конце. Цепь μ_m заканчивается гидрофобной последовательностью, которая, вероятно, и закрепляет ее на мембране. Эта последовательность замещается короткой гидрофильной последовательностью в μ_s; и такое замещение позволяет тяжелой цепи μ пройти через мембрану. Изменение С-конца совершается путем альтернативного сплайсинга 3'-конца ядерной РНК, как показано на РИС. 23.16.

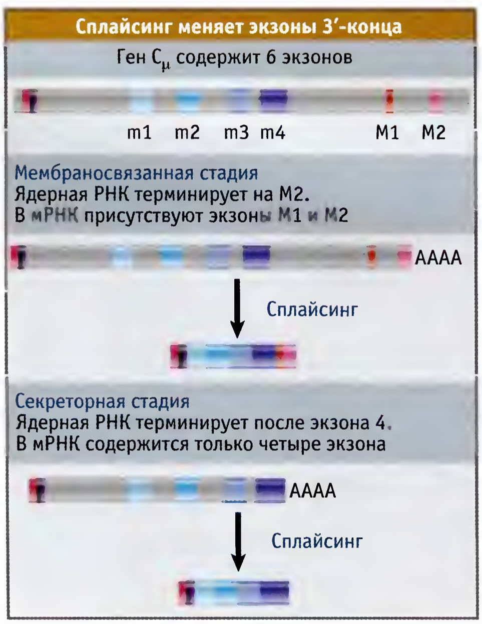


РИС. 23.16 3'-конец регулирует использование границ сплайсинга, таким образом, могут экспрессироваться альтернативные формы гена тяжелой цепи

На стадии существования мембраносвязанной формы IgM, РНК заканчивается после экзона M2, и константный участок образуется путем сплайсинга шести экзонов. Первые четыре экзона кодируют четыре домена константного участка. Последние два экзона, M1 и M2, кодируют гидрофобный С-концевой участок из 41 аминокислотного остатка и нетранслируемый трейлер. 5'-граница сплайсинга экзона 4 соединяется с 3'-границей сплайсинга в начале M1.

На секреторной стадии ядерная РНК оканчивается после экзона 4. 5'-граница сплайсинга этого экзона, которая соединялась с M1 в мембраносвязанной форме, сейчас игнорируется. Это позволяет экзону удлиниться на 20 кодонов.

Похожее преобразование с мембраносвязанной на секреторную форму обнаруживается и в случае с другими константными участками. Консервативность структуры экзонов говорит о том, что в данном случае действует тот же механизм.

23.12 Рекомбинация ДНК вызывает переключение классов иммуноглобулинов

Ключевые понятия:

- Иммуноглобулины подразделяют на пять классов в соответствии с типом константного участка тяжелой цепи.
- Переключение классов при смене C_H-участка происходит путем рекомбинации между S-районами с делецией всей последовательности между старым и новым C_H-участками.
- Может происходить множество успешных рекомбинаций для переключения классов.

Класс иммуноглобулина определяется типом C_H-участка. На РИС. 23.17 представлены пять классов Ig. IgM (иммуноглобулин, который В-клетка производит первым) и IgG (самый распространенный иммуноглобулин) сильнее других способны активировать комплемент, что ведет к разрушению внедрившейся клетки. IgA обнаруживается в выделяемых жидкостях (таких как слюна), а IgE участвует в аллергических реакциях и защите от паразитов.

Все лимфоциты начинают продуктивную жизнь как незрелые клетки, производящие IgM. Клетки, экспрессирующие IgM, имеют зародышевую организацию C_H-кластера генных сегментов, показанную на рис. 23.10. V-D-J реакции соединения запускают экспрессию C_μ-генного сегмента. Обычно лимфоцит производит один класс иммуноглобулинов в каждый момент времени, но класс может меняться в процессе клеточной дифференцировки. Такое изменение экспрессии называют **переключением классов**. Оно совершается путем замещения экспресси-

Доля	5%	1%	80%	14%	<1%
Эффекторная функция	Активируют комплемент	Отвечают за толерантность	Активируют комплемент	Обнаруживаются в секреторных жидкостях	Аллергические реакции

Рис. 23.17 Тип и функция иммуноглобулина определяются тяжелой цепью. J – это соединительный белок в IgM; все остальные типы Ig являются тетрамерами

рующегося C_H-участка. Факторы окружающей среды могут стимулировать переключение; например ростовой фактор TGFβ вызывает переключение с C_μ на C_α.

Переключение затрагивает только генный сегмент C_H; при этом продолжает экспрессироваться тот же V_H-генный сегмент. Таким образом, V_H-генный сегмент может последовательно экспрессироваться в комбинации с более чем одним C_H-генным сегментом. Та же самая легкая цепь продолжает экспрессироваться в клеточной линии дифференцировки. Переключение классов, следовательно, позволяет изменить тип эффекторного ответа (опосредуемого C_H-участком), сохраняя при этом способность распознавать тот же самый антиген (что опосредуется V-участком).

Изменения в экспрессии C_H-генного сегмента производятся двумя способами. Чаще всего это происходит путем дальнейшей рекомбинации ДНК (при использовании других систем, чем в случае соединения V-D-J, эти системы способны действовать позже в развитии B-клеток). Другой способ изменить класс действует на уровне процессинга РНК, но чаще этот способ используется для изменения C-концевой последовательности C_H-участка (см. разд. 23.11 «Процессинг РНК может изменить раннюю экспрессию тяжелой цепи»).

Клетки, экспрессирующие нижележащие C_H-генные сегменты, несут делеции C_μ и других генных сегментов, предшествующих экспрессируемому C_H-участку. Переключение классов осуществляется путем рекомбинации

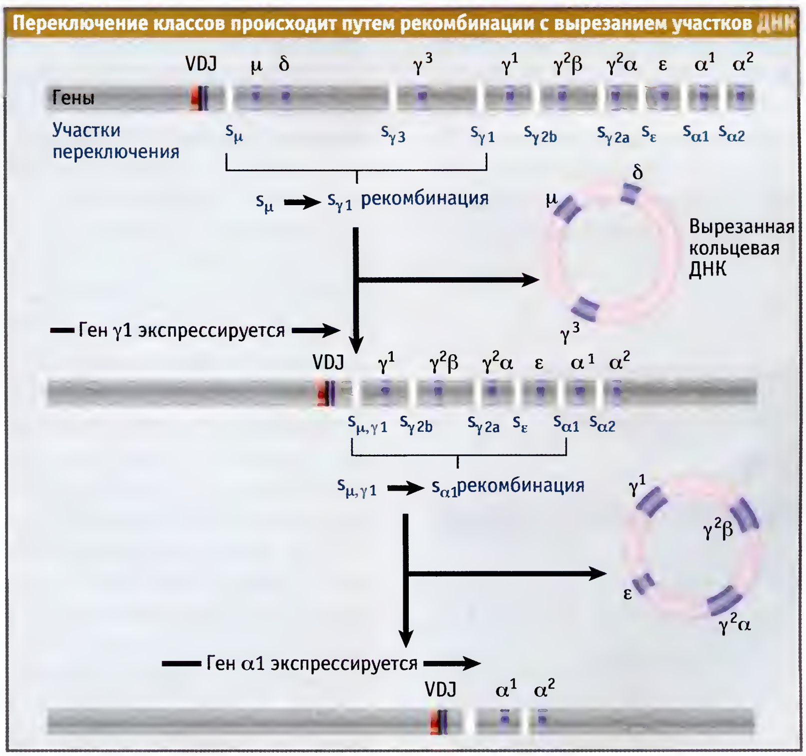


Рис. 23.18 Переключение классов генов тяжелых цепей может осуществляться путем рекомбинации между участками переключения (S), с делецией материала между рекомбинирующими S-сайтами. Возможны последовательные каскадные переключения

для сочленения нового C_H-генного сегмента с экспрессируемым участком V-D-J. Последовательности переключенных блоков V-D-J-C_H показывают, что сайты переключения лежат выше самих C_H-генных сегментов. Сайты переключения называют **S-участками**. На РИС. 23.18 показаны два последовательных переключения.

В первом переключении за экспрессией C_μ следует экспрессия $C_{\gamma1}$. $C_{\gamma1}$-генный сегмент переносится в положение для экспрессии путем рекомбинации между сайтами S_μ и $S_{\gamma1}$. Сайт S_μ лежит между V-D-J и C_μ-генным сегментом. Сайт $S_{\gamma1}$ расположен выше $C_{\gamma1}$-генного сегмента. Последовательность ДНК между двумя сайтами переключения вырезается в виде кольцевой молекулы.

Линейная делеционная модель налагает ограничения на локус тяжелых цепей: *После переключения классов невозможно экспрессировать любой из C_H-генных сегментов, которые располагались между C_μ- и новым C_H-генным сегментом.* На примере, показанном на рис. 23.18, клетки, экспрессирующие $C_{\gamma1}$, должны быть неспособны экспрессировать $C_{\gamma3}$ (который был удален.)

Однако сохраняется возможность выполнить последующее переключение и экспрессировать любой C_H-генный сегмент *ниже* уже экспрессированного гена. На рис. 23.18 показано второе переключение к экспрессии C_α, происходящее за счет рекомбинации между участком переключения $S_{\alpha1}$ и участком $S_{\mu,\gamma1}$, образованном в результате первого переключения.

Мы предполагаем, что все C_H-генные сегменты имеют S-участки выше кодирующей последовательности. Неизвестно, существуют ли какие-то ограничения на использование S-участков. Последовательные переключения действительно происходят, но мы не знаем, являются ли они факультативным или обязательным путем для перехода к более дальним C_H-генным сегментам. Интересно было бы узнать, может ли IgM переключаться напрямую на *любой* другой класс. S-участки лежат внутри интронов, предшествующих C_H-кодирующим участкам; поэтому переключение не изменяет рамку считывания.

23.13 Переключение происходит путем необычной реакции рекомбинации

Ключевые понятия:

- Переключение происходит с помощью двухцепочечного разрыва и соединения негомологичных частей.
- Характерной особенностью района переключения является наличие инвертированных повторов.
- Для переключения необходима активация промоторов выше сайта переключения.

Известно, что сайты переключения неоднозначны, поскольку доказано, что разные клетки, экспрессирующие одни и те же C_H-генные сегменты, рекомбинируют в разных местах. Участки переключения имеют различную длину (как определили по ограничениям сайтов, участвующих в рекомбинации) от 1 до 10 тпн. Они содержат группы коротких инвертированных повторов, в которых повторяющаяся единица может варьировать по длине от 10 до 80 нуклеотидов. Последовательность участка переключения, по-видимому, неважна; важно лишь наличие инвертированных повторов.

S-участок обычно располагается примерно в 2 тпн выше C_H-генного сегмента. Материал между сайтами переключения в результате реакции вырезается в виде кольцевой молекулы ДНК. Для фазы объединения в реакции VDJ-рекомбинации необходимы два белка: Ku и ДНК-PKcs (кроме того, они нужны и для негомологичного соединения концов, NHEJ). Этот факт говорит о том, что реакция соединения может использовать NHEJ путь. В целом это означает, что реакция происходит путем двухцепочечного разрыва, сопровождаемого воссоединением разрезанных концов.

Модель образования двухцепочечного разрыва имеет следующие основные моменты:

- нужна транскрипция S-участка;
- необходимы инвертированные повторы; и
- разрыв может произойти во многих местах внутри S-участка.

На РИС. 23.19 показаны стадии реакции переключения классов. Промотор (I) находится непосредственно перед каждым участком переключения. Для переключения этот промотор должен быть транскрибирован. Промотор может отвечать на действия активаторов, которые, в свою очередь, отвечают на условия среды (например, на стимуляцию цитокинами), и эта система, таким образом, представляет собой механизм для регуляции переключения. Следовательно, первая стадия переключения — это активация I-промоторов перед каждым участком, который будет участвовать в реакции. При активации с этих промоторов образуются нетранслируемые (стерильные) транскрипты, которые сплайсируются для соединения I-участка с соответствующим участком тяжелой цепи.

Ключевым событием на пути к пониманию механизма переключения было открытие роли фермента AID (активационно-индуцируемой цитидиндезаминазы). В отсутствие AID переключение классов блокируется до стадии образования одноцепочечного разрыва. Соматический мутагенез также блокируется, что указывает на интересную связь двух важных процессов генерации иммунологического разнообразия (см. разд. 23.15 «Цитидин-дезаминаза и урацил-гликозилаза вызывают соматический мутагенез»).

AID экспрессируется только на определенной стадии дифференцировки В-лимфоцитов, привязывая процессы переключения классов и соматического мутагенеза к этой стадии. AID — один из членов класса ферментов, воздействующих на РНК для замены цитидина на уридин (см. разд. 27.10 «Редактирование РНК касается конкретных азотистых оснований»). Однако AID имеет другую специфичность и действует на одноцепочечную РНК.

Для переключения классов и соматического мутагенеза также необходим дополнительный фермент. Это UNG,

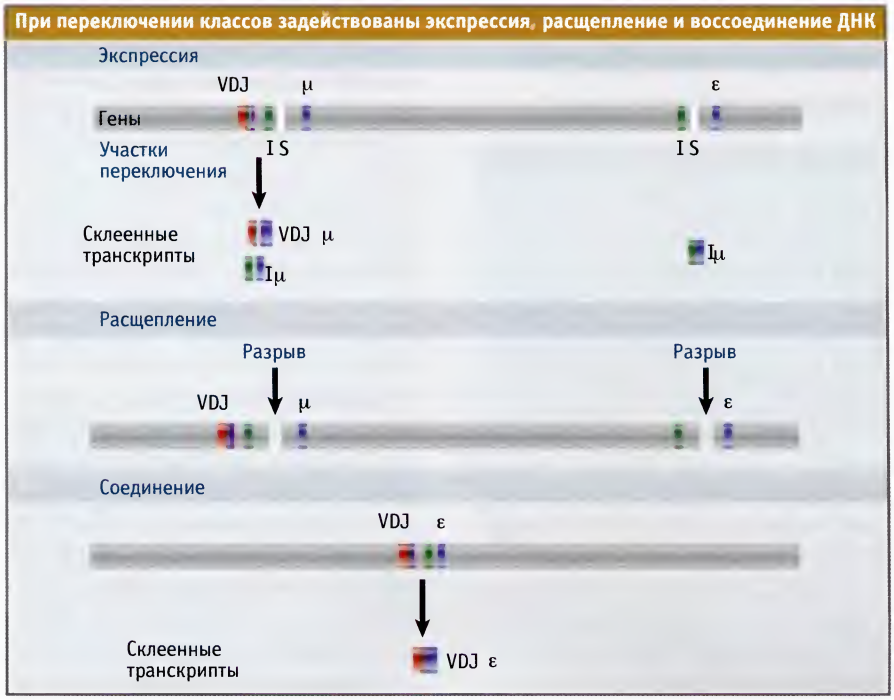


РИС. 23.19 Переключение классов происходит через отдельные стадии. I-промоторы инициируют транскрипцию с образованием нетранслируемых (стерильных) транскриптов. Переключаемые участки вырезаются. В местах вырезанных участков происходит соединение

урацил-ДНК-гликозилаза, удаляющий урацил, который AID образует путем дезаминирования цитидина. Дефицитные по UNG мыши имеют сниженный в 10 раз уровень переключения классов. Это предполагает модель, при которой последовательные действия AID и UNG создают такие сайты в ДНК, из которых удалено основание. Такое действие приводит к различным последствиям в переключении классов и в системе соматического мутагенеза.

Источник одноцепочечной ДНК-мишени для AID образуется путем стерильной транскрипции, наиболее вероятно, из-за экспонирования нематричной цепи ДНК, вытесняемой когда другая цепь используется в качестве матрицы для синтеза РНК. Эта модель поддерживается наблюдением, что AID выбирает в качестве мишени предпочтительно нематричную цепь.

Для переключения классов эти сайты превращаются в разрывы нуклеотидной цепочки, что обеспечивает вырезание участков, показанное на рис. 23.19. Разрезанные концы соединяются путем NHEJ, который представляет собой систему репарации, действующую на двухцепочечных разрывах ДНК (см. разд. 20.12 «Неспециализированная система репарации двухцепочечных разрывов»). Пока неизвестно, как участок с удаленным основанием превращается в двухцепочечный разрыв. Возможно, в этом задействована система MSH (magnesium silicate hydrous), участвующая в репарации неспаренных оснований в ДНК (поскольку мутации в гене *MSH2* снижают уровень переключения классов).

Одна из неожиданных особенностей процесса — это участие инвертированных повторов. Существует вероятность того, что шпильки образуются путем взаимодействия между инвертированными повторами на вытесняемой нематричной цепи, как показано на РИС. 23.20. При сопряжении с образованием участка с удаленным основанием на этой цепи данный процесс может приводить к разрыву.

Два очень важных вопроса, которые остались без ответа, состоят в следующем: как описанная система нацеливается на соответствующие участки в локусе тяжелых цепей и как происходит контроль за использованием сайтов переключения?

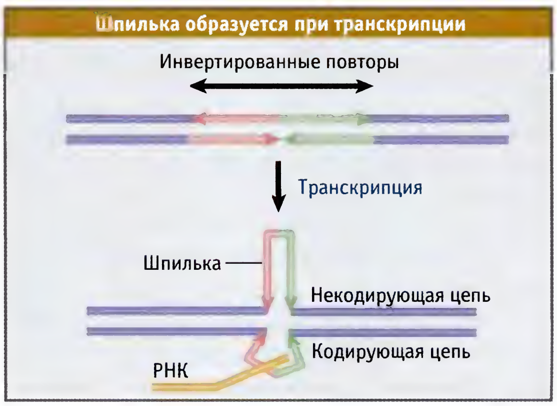


РИС. 23.20 В момент разделения цепей ДНК при транскрипции одна из цепей может образовывать шпильку (при наличии палиндромной последовательности)

23.14 Соматические мутации создают дополнительное многообразие антител у мышей и человека

Ключевые понятия:

- Последовательность V-районов активных генов иммуноглобулинов отличается от этого участка в клетках зародышевого типа из-за соматических мутаций.
- Соматические мутации представляют собой замену конкретных нуклеотидов.
- Сайты мутаций концентрируются в области антиген-связывающего участка.
- Внесение соматических мутаций зависит от работы энхансера, который активирует транскрипцию Ig-локуса.

При сравнении последовательностей экспрессируемых генов иммуноглобулинов и соответствующих V-генных сегментов зародышевой линии видно, что в экспрессируемой популяции появляются новые последовательности. Некоторая часть этого дополнительного разнообразия образуется в результате изменения последовательностей в V-J- или V-D-J-соединениях, происходящего во время рекомбинации. Другие изменения, однако, наблюдаются выше этих участков внутри вариабельных доменов.

После того как путем реорганизации был образован функциональный ген иммуноглобулина, изменения в V-генных последовательностях происходят с помощью двух механизмов. У человека и мыши таким механизмом является индукция **соматических мутаций** в отдельных точках внутри гена, в особенности в активных лимфоцитах. Этот процесс иногда называют **гипермутацией**. У кур, кролика и свиньи другой механизм использует генную конверсию для замены сегмента экспрессируемого V-гена на соответствующую последовательность из другого V-гена (см. разд. 23.16 «Иммуноглобулины птиц образуются из псевдогенов»).

Зонд, представляющий собой участок экспрессируемого V-гена, можно использовать для определения всех соответствующих участков в зародышевой линии. Их последовательности должны показать весь спектр таких сегментов, которые могут использоваться в организме. Любой экспрессируемый ген, последовательность которого отличается, образовался в результате соматических изменений.

Одна из трудностей состоит в том, чтобы доказать, что были обнаружены действительно все потенциальные источники V-генных сегментов зародышевой линии. Эта проблема не представляет трудности в случае мыши, у которой достаточно простая система λ-цепей. Исследование нескольких миеломных линий, производящих λ_1-цепи, показало, что часть из них имеют последовательность одного и того же генного сегмента зародышевой линии. *Другая часть несет новые последовательности, которые должны были образоваться путем мутаций в зародышевом генном сегменте.*

Для определения частоты соматических мутаций необходимо исследовать большое число клеток, в которых экспрессируются одинаковые V-генные сегменты. На практике для определения такой группы характеризуют иммуноглобулины серии клеток, проявляющих иммунный ответ к определенному антигену.

Для этой цели используют эпитопы к малым молекулам — гаптенам, элементарная структура которых вызывает стабильный ответ в отличие от многих белков, различные части которых вызывают образование различных антител. Гаптены конъюгируют с нереактивным белком для образования антигена. Клетки получают, иммунизируя мышь таким антигеном, выделяют реактивные лимфоциты и иногда сливают эти лимфоциты с клетками миеломы (иммортализованными опухолевыми клетки) для образования **гибридом**, которые продолжают экспрессировать нужные антитела неограниченно долгое время.

Приведем один из примеров: 10 из 19 различных клеточных линий, производящих антитела против гаптена фосфорилхолина, имеют одну и ту же V_H-последовательность. Эта последовательность имеет зародышевый V-генный сегмент T15, один из четырех родственных V_H-генов. Другие девять экспрессируемых генных сегментов отличаются друг от друга и от всех четырех зародышевых членов семейства. Они были более похожи на T15 зародышевую последовательность, чем на какую-либо иную, и их фланкирующие последовательности были такими же, как и вокруг T15. Этот факт говорит о том, что они возникли из T15 путем соматического мутагенеза.

На РИС. 23.21 показано, что изменения последовательности локализованы вокруг V-генного сегмента, располагаясь на участке от ~150 пн выше V-генного промотора и до ~1,5 тпн. Изменения представляют собой замены отдельных пар нуклеотидов. Обычно это от ~3 до ~15 замен, что соответствует <10 заменам аминокислот на белок. Замены концентрируются в антиген-связывающем сайте (создавая, таким образом, максимум разнообразия для распознавания новых антигенов). Только некоторые из мутаций влияют на аминокислотную последовательность, поскольку часть из них оказывается в третьем положении кодонов, а также в нетранслируемых участках.

Присутствие довольно большой части незначащих мутаций говорит о том, что соматические мутации про-



РИС. 23.21 Соматические мутации происходят в районе, окружающем V-сегмент, и распространяются на соединенные VDJ-сегменты

исходят более или менее случайным образом в районе, включающем V-генный сегмент и его окружение. Для некоторых мутаций существует тенденция повторяться во многих случаях. Они могут представлять собой горячие точки, возникающие из-за каких-то внутренних предпочтений в системе.

Соматические мутации происходят во время клональной пролиферации, со скоростью примерно ~10^{-3} на пн на поколение клеток. Примерно половина дочерних клеток несет мутацию; в результате клетки, экспрессирующие мутантные антитела, составляют бóльшую часть клона.

Во многих случаях единственное семейство V-генных сегментов постоянно используется для ответа на определенный антиген. При узнавании антигена V-участок с наибольшей аффинностью (предположительно) обеспечивает отправную точку. Соматические мутации затем увеличивают спектр разнообразия. Случайные мутации имеют непредсказуемые эффекты на функцию белка; некоторые инактивируют белок, в то время как другие обеспечивают высокую специфичность к определенному антигену. Пропорция и эффективность лимфоцитов, отвечающих на антиген, увеличиваются путем селекции в популяции с отбором клеток, несущих антитела, мутации в которых увеличили аффинность к антигену.

23.15 Цитидиндезаминаза и урацилгликозилаза вызывают соматический мутагенез

Ключевые понятия:

- Цитидиндезаминаза необходима как для соматического мутагенеза, так и для переключения классов иммуноглобулинов.
- Активность урацил-ДНК-гликозилазы влияет на структуру соматических мутаций.
- Последовательное действие этих ферментов может вызвать гипермутацию.

Процесс возникновения соматических мутаций предъявляет такие же требования, как и переключение классов (см. разд. 23.13 «Переключение происходит путем необычной реакции рекомбинации»):

- целевой участок должен быть транскрибирован (в этом случае необходим энхансер, активирующий транскрипцию в каждом Ig-локусе);
- необходимы ферменты AID и UNG, и
- MSH-система репарации неспаренных оснований также необходима.

Модель, при которой удаление дезаминированного основания приводит к соматическим мутациям, предложена на основе экспериментов, описанных на РИС. 23.22. Когда

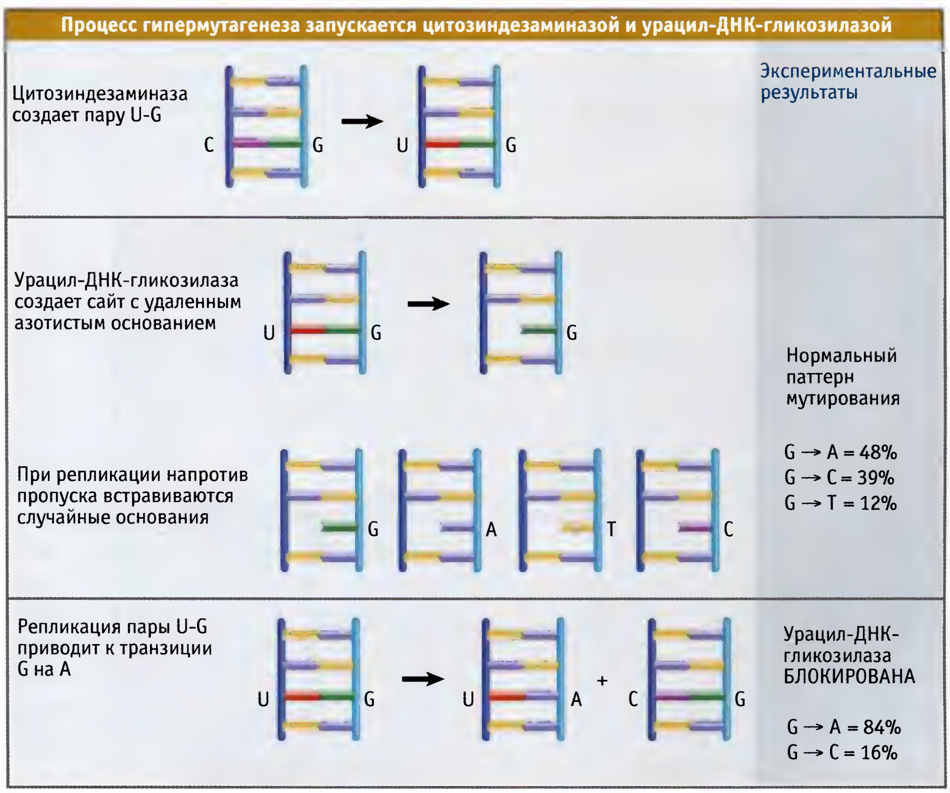


РИС. 23.22 Когда действие цитидиндезаминазы (вверху) сопровождается действием урацил-ДНК-гликозилазы, образуется сайт с удаленным азотистым основанием. При репликации в этот сайт (в дочерней цепи) с равной вероятностью могут встроиться все четыре основания (в центре). Если же урацил из ДНК не удаляется, то репликация с его участием приводит к транзиции C-G на A-T

AD дезаминирует цитозин, образуется урацил, который затем удаляется из ДНК ферментом UNG. Обычно все возможные замены происходят в сайте с удаленным азотистым основанием. Если действие урацил-ДНК-гликозилазы блокировано, то мы увидим другой результат. Если урацил из ДНК не был удален, при репликации он спарится с аденином. В итоге исходная C-G пара заменится на A-T пару. Урацил-ДНК-гликозилаза может быть блокирована путем введения в клетки гена, кодирующего белок, ингибирующий этот фермент. (Этот ген — компонент бактериофага PSB-2, чей геном содержит урацил, таким образом, ему необходимо блокировать этот фермент при фаговой инфекции.) Когда такой ген вводят в лимфоцитарную клеточную линию, происходит сильное изменение паттерна мутаций, причем почти все мутации представляют собой предполагаемые транзиции C-G на A-T.

Таким образом, главное событие в образовании случайных мутаций — это создание сайта с удаленным азотистым основанием. Затем MSH-система репарации производит вырезание и замену участка ДНК, содержащего повреждение. Мутации могут возникать из-за ошибок ДНК-полимеразы. Другим источником мутаций служат события, в которых образуется так много сайтов с отсутствующим азотистым основанием, что система репарации не справляется с таким количеством нарушений. При последующей репликации это может приводить к случайному внедрению оснований напротив сайтов с отсутствующим основанием. Пока неизвестно, что именно ограничивает действие этой системы целевым районом гипермутирования.

Разница между системами наблюдается в конце процессов, когда в случае с переключением классов вносятся двухцепочечные разрывы, а в процессе соматического мутагенеза образуются индивидуальные точечные мутации. Пока неизвестно, где эти системы начинают различаться. Возможно, это происходит при внесении разрывов в сайты с удаленным азотистым основанием при переключении классов, в то время как при соматическом мутагенезе эти сайты подлежат репарации. Также, возможно, разрывы создаются в обоих случаях, но репарируются с ошибками при соматическом мутагенезе.

23.16 Иммуноглобулины птиц образуются из псевдогенов

Ключевые понятия:

- Ген иммуноглобулина у кур формируется путем копирования последовательности одного из 25 псевдогенов в единственный активный локус V-гена.

Иммунная система кур — пример системы, встречающейся еще у ряда организмов (кролика, коровы, свиньи) и использующей многообразие, закодированное в геноме. Схожий механизм используется как единственным локусом легких цепей (λ-типа), так и локусом H-цепей. Организация λ-локуса показана на РИС. 23.23. В данном локусе есть один функциональный V-генный сегмент, J-сегмент и C-генный сегмент. Выше функционального $V_{\lambda 1}$-генного сегмента лежат 25 V_{λ} псевдогенов (в обеих ориентациях). Их классифицируют как псевдогены, поскольку в них либо делетирован кодирующий сегмент с одного или с обоих концов, либо

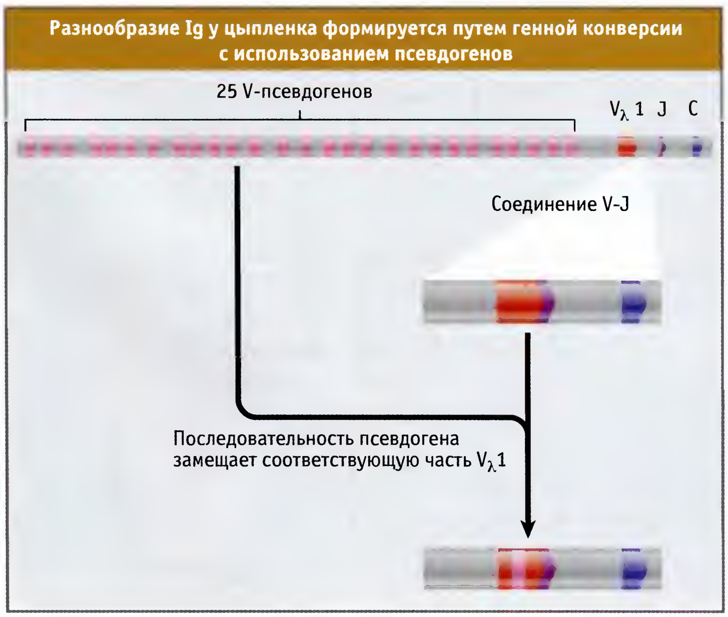


РИС. 23.23 Локус легких цепей λ у кур несет 25 V-псевдогенов выше единственного функционального V-J-C участка. Однако последовательности псевдогенов обнаруживаются в активных реорганизованных V-J-C генах

отсутствуют необходимые сигналы для рекомбинации, либо встречаются сразу обе эти особенности. Принадлежность к псевдогенам подтверждается также тем фактом, что только $V_{\lambda 1}$-генный сегмент рекомбинирует с J-C_λ-генным сегментом.

Однако последовательности активных реорганизованных V_λ-J-C_λ-генных сегментов очень разнообразны. Реорганизованный ген имеет один или более участок, в котором встречаются кластеры измененных последовательностей. Последовательность, идентичную новой последовательности, почти всегда можно найти в одном из псевдогенов (которые сами по себе остаются неизменными). Необычные последовательности, отсутствующие среди псевдогенов, всегда представляют собой изменения в местах соединения между исходной и измененной последовательностью.

Таким образом, здесь используется новый механизм получения многообразия. Последовательности из псевдогенов длиной от 10 до 120 пн замещаются в активной $V_{\lambda 1}$-области путем генной конверсии. Немодифицированные $V_{\lambda 1}$-последовательности не экспрессируются даже на ранних стадиях иммунного ответа. Успешное конверсионное событие, по-видимому, происходит каждые десять–двадцать клеточных делений на каждую реорганизованную $V_{\lambda 1}$-последовательность. В конце периода иммунологического созревания реорганизованная $V_{\lambda 1}$-последовательность имеет от четырех до шести преобразованных сегментов (покрывающих весь участок), произошедших из различных донорных псевдогенов. Если участвуют все псевдогены, получается около $2{,}5 \times 10^8$ возможных комбинаций!

Ферментативные основы для копирования последовательностей псевдогенов в экспрессионный локус зависят от ферментов, участвующих в рекомбинации, и схожи с механизмом соматического гипермутагенеза, создающим многообразие у человека и мыши. Некоторые из генов, участвующих в рекомбинации, необходимы для процесса генной конверсии; например, конверсия прекращается при делеции *RAD54*. Делеция ряда генов, участвующих в рекомбинации (*XRCC2*, *XRCC3* и *RAD51B*), приводит к другому интересному эффекту: в экспрессионном локусе V-гена возникают соматические мутации. Частота соматических мутаций примерно в 10 раз выше, чем при обычной скорости генной конверсии.

Описанные факты показывают, что соматические мутации у кур отсутствуют не из-за неполноценности ферментативной системы (имеющейся у мыши и человека). Наиболее вероятное объяснение связи между рекомбинацией и соматическим мутагенезом состоит в том, что нерепарированные разрывы индуцируют мутации. Причина того, почему соматические мутации происходят у человека и мыши, но не у кур, может заключаться в деталях действия системы репарации, которая работает над разрывами в этих локусах. Репарация более эффективна у кур, поэтому ген может быть репарирован путем генной конверсии раньше, чем будут индуцированы мутации.

23.17 Быстрый вторичный ответ возможен благодаря В-клеткам памяти

Ключевые понятия:

- Первичный ответ на антиген формируется благодаря В-клеткам, погибающим после ответа.
- Возникают неактивные В-клетки памяти, имеющие специфичность к тому же самому антигену.
- Повторная встреча с антигеном запускает вторичный ответ, и клетки памяти быстро активируются.

Теперь мы готовы подвести итог взаимодействия между производством высокоаффинных антител и дифференцировкой В-клеток. На РИС. 23.24 показано, что В-клетки

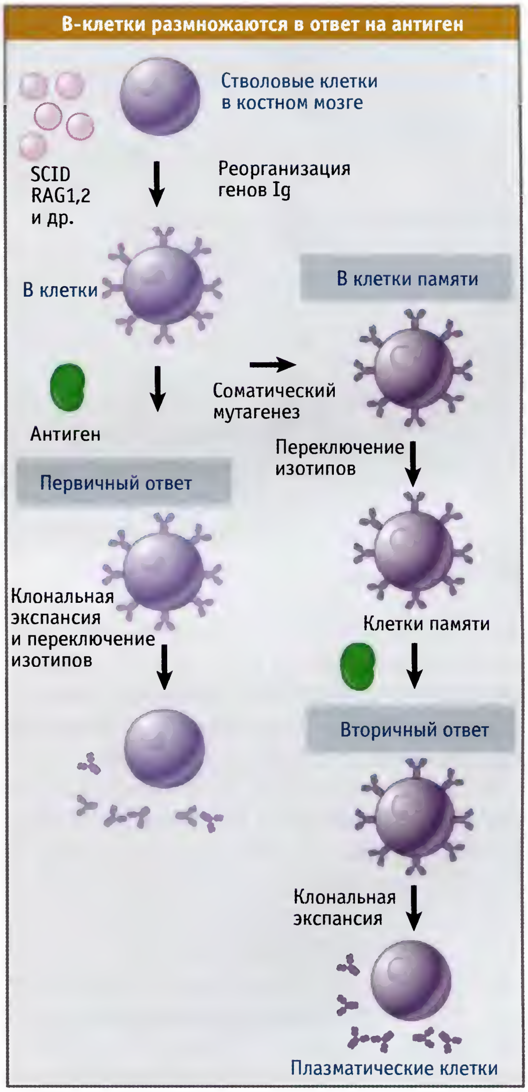


РИС. 23.24 Приобретенный иммунитет обеспечивается дифференцировкой В-клеток. Пре-В-клетки превращаются в В-клетки путем реорганизации генов иммуноглобулинов. Первичная встреча с антигеном вызывает как первичный ответ, так и хранение клеток памяти. Последующая встреча с антигеном вызывает вторичный ответ клеток памяти

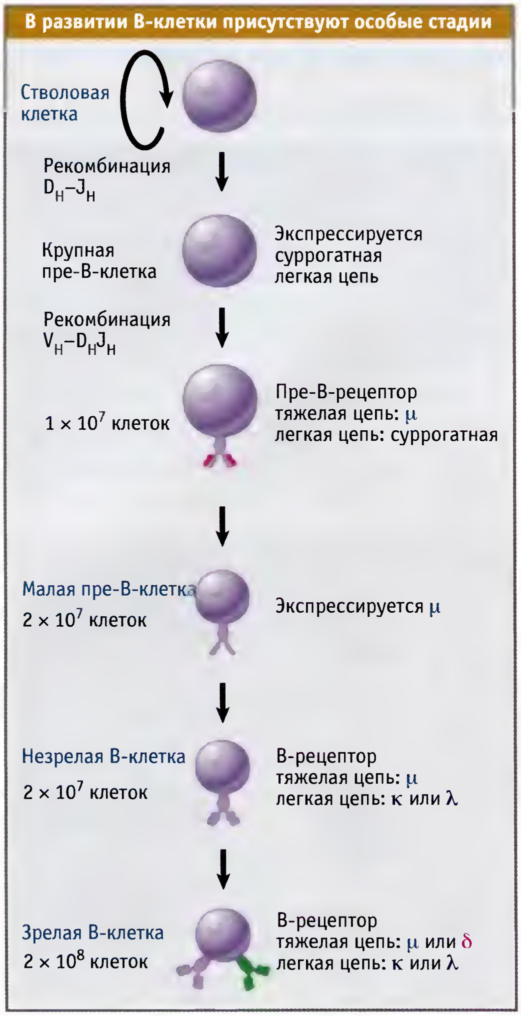


РИС. 23.25 В-клетка в своем развитии проходит через ряд последовательных стадий

происходят из самообновляющейся популяции стволовых клеток костного мозга. Созревание В-клеток зависит от реорганизации генов Ig, для которой необходимо функционирование генов *SCID* и *RAG1,2* (и других). Если реорганизация генов блокируется, зрелые В-клетки не образуются. Антитела В-клеток обладают специфичностями, определяемыми особыми комбинациями V(D)J-участков и дополнительными нуклеотидами, вводимыми во время процесса соединения.

Встреча с антигеном запускает два аспекта иммунного ответа. **Первичный иммунный ответ** проявляется путем клональной экспансии В-клеток, отвечающих на антиген. В результате этого образуется большое число плазматических клеток, специфичных к антигену; для выработки сответствующего типа эффекторного ответа происходит переключение изотипа. Популяция клеток, участвующих в первичном ответе, обречена: эти клетки не живут дольше самого первичного ответа.

Вторичный иммунный ответ возможен благодаря существованию **В-клеток памяти**. Соматический мутагенез создает В-клетки с повышенной аффинностью к антигену. Эти клетки не запускают немедленный иммунный ответ, хотя и могут подвергаться переключению изотипов для выбора других форм C_H-участка. Они хранятся как клетки памяти с соответствующей специфичностью и эффекторным типом ответа, но являются неактивными. Клетки памяти активируются при повторном контакте с тем же антигеном. Они подготовлены к встрече с данным антигеном и в результате позволяют вторичному ответу развиться очень быстро путем клональной экспансии. В ходе вторичного ответа не происходит ни дальнейшего соматического мутагенеза, ни переключения изотипов.

Процессы, изображенные на рис. 23.24, показывают развитие приобретенного иммунитета, то есть ответа на определенный антиген. Кроме описанных, имеется отдельное множество В-клеток, называемых Ly-1-клетками. Эти клетки подверглись процессу реорганизации V-гена и, вероятно, отобраны для экспрессии определенного пула антител. Они не подвергаются соматическому мутагенезу или ответу памяти, а участвуют в естественном иммунитете (то есть во врожденной способности реагировать на некоторые антигены).

Более детально процесс развития В-клеток показан на **РИС. 23.25**. Первый этап — это рекомбинация между D-сегментом и J-сегментом тяжелой цепи μ. За ней следует рекомбинация V-D, создающая тяжелую цепь μ. Может происходить несколько рекомбинационных событий, включающих последовательность продуктивных и непродуктивных реорганизаций (см. рис. 23.23). Эти клетки выделяют белок, напоминающий λ-цепь и называемый суррогатной легкой цепью, который экспрессируется на поверхности и ассоциирует с тяжелой цепью μ, форми-

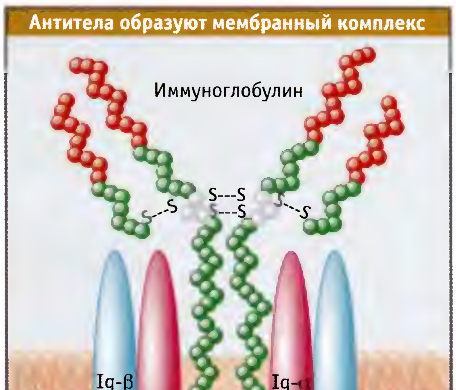


РИС. 23.26 В-клеточный антигенный рецептор состоит из тетрамерного иммуноглобулина (H_2L_2), связанного с двумя копиями проводящего сигнал гетеродимера (Igαβ)

руя пре-В-рецептор. Он похож на иммуноглобулинный комплекс, но действует иным образом.

Выработка μ-цепи подавляет синтез цепи SL, и клетки делятся, образуя небольшие пре-В-клетки. Затем легкая цепь экспрессируется, и действующий иммуноглобулин появляется на поверхности незрелой В-клетки. Происходит еще одно деление, и по мере взросления В-клетки к экспрессии μ-цепи добавляется экспрессия тяжелой цепи δ.

Иммуноглобулины выполняют свои функции как при секреции В-клетками, так и при экспрессии на поверхности. На РИС. 23.26 показано, что активный комплекс на поверхности клетки, называемый **рецептором В-клетки (BCR)**, состоит из иммуноглобулина, ассоциированного с трансмембранными белками Igα и Igβ. Они обеспечивают сигнальные компоненты, включающие межклеточные пути в ответ на связывание антигена с антителом. На активирование BCR также влияет взаимодействие с другими рецепторами, например участие во взаимодействии В-клеток, активированных антигеном, и Т-хелперов.

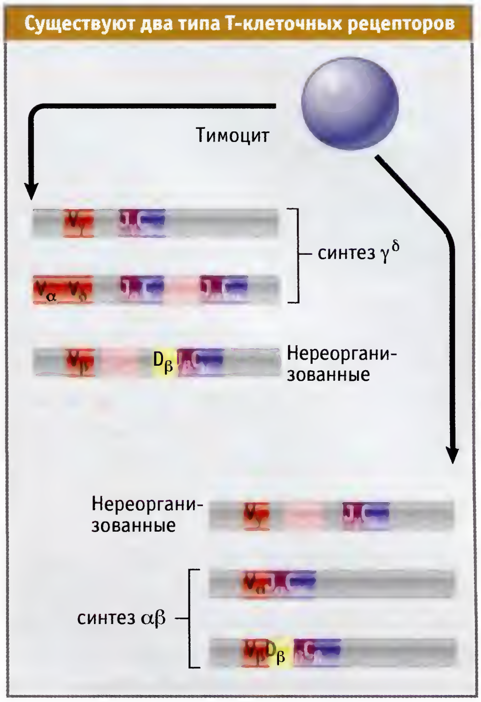


РИС. 23.27 γδ-рецептор синтезируется на ранней стадии развития Т-клеток. TCR αβ синтезируется позже и отвечает за «классический» клеточно-опосредованный иммунитет, в котором целевой антиген и хозяйский антиген гистосовместимости распознаются совместно

23.18 Рецепторы Т-клеток схожи с иммуноглобулинами

Ключевые понятия:

- Для образования одного из двух типов рецепторов Т-клеток Т-клетки используют аналогичный В-клеточному механизм V-(D)J-соединения.
- Т-рецепторы αβ обнаруживаются на поверхности более 95% Т-лимфоцитов; Т-рецепторы γδ — на поверхности менее 5% Т-лимфоцитов.

Линия лимфоцитов представляет собой пример эволюционного оппортунизма: В-клетки и Т-клетки используют схожие процессы для выработки белков с вариабельным участком (обеспечивающим существенное разнообразие) и более ограниченными константными участками, отвечающими за небольшой набор эффекторных функций. Т-клетки производят один из двух типов рецепторов. Разные рецепторы синтезируются на разных этапах развития Т-клетки, как показано на РИС. 23.27.

Рецептор γδ обнаруживается менее чем на 5% Т-лимфоцитов. Он синтезируется только на ранней стадии развития Т-клетки. У мышей это единственный рецептор, обнаруживаемый на сроках беременности меньше 15 дней, но к моменту рождения (на 20-й день) он почти не встречается.

Рецептор αβ присутствует более чем на 95% лимфоцитов. Он синтезируется в более поздней фазе развития Т-клеток, чем γδ. У мышей он начинает обнаруживаться на 15–17-й день беременности, а к моменту рождения становится доминирующим. Данный рецептор синтезируется отдельной линией клеток, не связанной с клетками, занятыми в выработке γδ, и использует собственные события реорганизации.

Подобно иммуноглобулинам, TCR должны распознавать внешний антиген заранее неизвестной структуры. Проблема распознавания антигенов у Т-клеток решается так же, как у В-клеток, а строение рецепторных генов Т-клетки по использованию вариабельных и константных участков напоминает гены иммуноглобулинов. *Каждый локус организован подобно генам иммуноглобулинов, с отдельными сегментами, соединяемыми с помощью реакций рекомбинации, уникальных для данного лимфоцита.* Те же самые компоненты обнаруживаются и у трех семейств Ig.

Строение белков TCR напоминает строение иммуноглобулинов. Внутренняя организация V-последовательностей у белков Ig и TCR общая. С-участок TCR соответствует константным участкам Ig и включает единственный константный домен, за которым следуют трансмембранная и цитоплазматическая части. Экзон-интронная структура соответствует функции белка.

Сходство генов TCR и Ig поразительно. Как показано на РИС. 23.28, строение TCR α напоминает строение Ig κ, где V-сегменты гена отделены от кластера J-сегментов, предшествующего единственному С-сегменту. Строение этого локуса сходно у человека и мыши, некоторые различия имеются лишь в числе сегментов V_α и J_α. (Помимо α-сегментов, этот локус содержит и δ-сегменты, которые мы вскоре обсудим.)



РИС. 23.28 **В локусе TCRα у человека находятся рассеянные α- и δ-сегменты. V_δ-сегмент расположен внутри кластера Vα. D-J-C_δ-сегменты расположены между V-генными сегментами и J-C_α-сегментами. У мыши этот локус организован сходным образом, но содержит большее число V_δ-сегментов**

Компоненты Т-рецепторов β сходны с компонентами IgH. РИС. 23.29 показывает, что их организация различна, и V-сегменты отделены от двух кластеров, каждый из которых включает D-сегмент, несколько J-сегментов и С-сегмент. Единственое различие человека и мыши и в этом случае заключается в числе V_β J_β.

Разнообразие создается теми же механизмами, что и у иммуноглобулинов. Врожденная дифференциация возникает в результате сочетания различных V, D, J и С-сегментов; дополнительная возникает в результате вставки новых последовательностей (в виде Р и N-нуклеотидов, см. рис. 23.14) в местах соединения этих компонентов. Некоторые цепи β включают два D-сегмента, возникающих за счет соединения D-D (управляемого соответствующими структурами нонамерных и гептамерных последовательностей). Различие между TCR и Ig состоит в отсутствии соматического мутагенеза у локусов. Измерения степени разнообразия показывают, что 10^{12} Т-клеток человека содержат 2,5 × 10^7 различных α-цепей, ассоциированных с 10^6 различных β-цепей.

Строение γ-локуса сходно с Ig λ, где V-сегменты гена отделены от последовательностей сегментов J-C. РИС. 23.30 демонстрирует сравнительно небольшое разнообразие этого локуса (примерно 8 функциональных V-сегментов). Его строение у человека и у мыши различно. Мыши имеют 3 функциональных J-C локуса, но некоторые сегменты обратно ориентированы. У человека на каждый С-сегмент приходится по нескольку J-сегментов.

δ-единицы кодируются сегментами, находящимися в α-локусе, как было показано на рисунке 23.28. Сегменты D_δ-D_δ-J_δ-C_δ расположены между V и J_α-C_α-сегментами. Оба D-сегмента могут быть включены в δ-цепь, образуя структуру VDDJ. Интересен вопрос природы V-сегментов, используемых в δ-перестановке. В активных TCR δ-цепях обнаруживается очень небольшое число V-последовательностей. У человека для этого используется всего один V-сегмент. У мыши обнаружены несколько сегментов V_δ, из которых некоторые специфичны для δ-перестановки, но некоторые обнаруживаются также и в α-перестановках. Причина специфичности выбора V-сегмента в α и δ перестановках неизвестна. Возможно, дело в том, что многие сегменты V_α могут сцепляться с сегментом DDJ_δ, но лишь немногие (определяемые как V_δ) могут создавать активные белки.

На данный момент мы обозначили V-сегменты в δ-цепях как V_δ, однако мы пока не можем однозначно утверждать, являются ли они уникальными для δ-перестановки. Рассеянное положение генов предполагает, что рецепторы TCR αβ и γδ являются взаимоисключающими в пределах одного аллеля, поскольку δ-локус полностью утрачивается при перестановке V_α-J_α.

Перестановки в локусе TCR, как и в гене иммуноглобулина, могут быть продуктивными или не продуктивными. β-локус демонстрирует аллельное исключение подобно иммуноглобулиновым локусам; перестановка прекращается, как только создан продуктивный аллель. α-локус может отличаться; случаи продолжающейся перестановки предполагают возможность того, что замена последовательностей V_α может продолжаться и после создания продуктивного аллеля.



РИС. 23.29 **В локусе TCR_β содержится множество V-генных сегментов, лежащих на участке около 500 тпн, расположенном ~280 тпн выше двух D-J-C-кластеров**

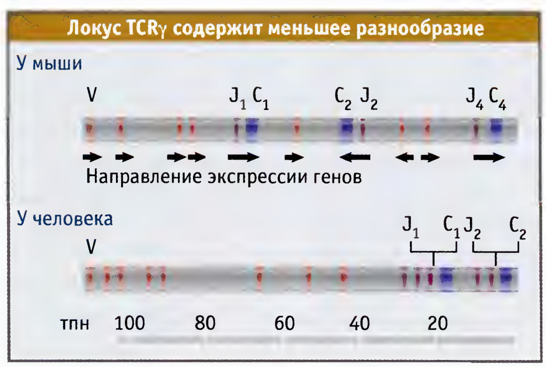


РИС. 23.30 **В локусе TCRγ содержится небольшое число функциональных V-генных сегментов (некоторые псевдогены не показаны), которые лежат выше локусов J-C**

23.19 Рецепторы Т-клеток действуют совместно с МНС

Ключевые понятия:

- Т-рецепторы распознают короткий пептид, расположенный в бороздке белка главного комплекса гистосовместимости на поверхности презентирующей клетки.

Т-клетки с αβ-рецепторами делят на несколько подтипов, имеющих разные функции (связанные со взаимодействием клеток, задействованных в иммунном ответе). Цитотоксические Т-клетки обладают способностью лизировать инфицированную целевую клетку. Т-хелперы помогают в разрушении чужеродных клеток и взаимодействии антиген-антитело (опосредованном В-клетками).

Основное различие между антителами В-клеток и рецепторами Т-клеток заключается в характере их взаимодействия с антигеном. Антитело распознает короткий участок антигена (эпитоп). Рецептор Т-клетки связывается с коротким (4–5 аминокислот) пептидом, который вырезается из антигена путем его расщепления. Этот пептидный фрагмент презентируется Т-клетке белком главного комплекса гистосовместимости (МНС). Таким образом, Т-клетка одновременно распознает внешний антиген и белок МНС, доставленный презентирующей клеткой, как было показано на рис. 23.2. В процессе участвуют как Т-хелперы, так и Т-киллеры, но требования к презентированию антигена у них различны; в разных случаях используются различные виды белка МНС (см. разд. 23.20 «Локус главного комплекса гистосовместимости кодирует множество генов иммунной системы»). Т-хелперам требуется, чтобы антиген был презентирован белком МНС класса II, а Т-киллерам — белком МНС класса I.

αβ-Рецепторы TCR отвечают за функцию хелперных Т-клеток в гуморальном иммунитете и функцию киллерных Т-клеток в клеточно-опосредованном иммунитете. Это накладывает на них обязанность распознавать как внешний антиген, так и собственый МНС-белок. Белок МНС связывает короткий пептид, полученный из внешнего антигена, после чего TCR распознает этот пептид в бороздке на поверхности МНС. Говорят, что МНС *презентирует* пептид рецептору TCR. (Пептид образуется в результате расщепления протеасомой белка-антигена на фрагменты длиной 8–10 остатков.) Конкретный TCR специфичен для определенного МНС и определенного внешнего антигена. Основа этой «двойной» способности — один из наиболее интересных нерешенных вопросов, касающихся αβ TCR.

Рекомбинация, создающая рабочие цепочки TCR, связана с развитием Т-лимфоцита (см. РИС. 23.31). Сначала происходит реорганизация, в результате которой появляется активная β-цепь TCR. Она связывается с нереорганизованой суррогатной α-цепью, называемой пре-TCRα. На этой стадии лимфоцит еще не экспрессировал на своей поверхности белок CD4 или CD8. Затем гетеродимер пре-TCR связывается с сигнальным комплексом CD3 (см. ниже). Сигналы с этого комплекса запускают процесс нескольких клеточных делений, во время которых α-цепи реорганизуются, а гены CD4 и CD8 начинают работать, превращая лимфоциты из $CD4^-CD8^-$ в $CD4^+CD8^+$. Этот этап развития носит название *β-селекции*. В ходе него образуются **ДП-тимоциты**.

Рекомбинация α-цепи продолжается в DP-тимоцитах. Процесс созревания продолжается как позитивной селекцией (для зрелых TCR-комплексов, способных связываться с лигандами), так и негативной селекцией (для комплексов, ошибочно взаимодействующих с собственными лигандами). Оба вида селекции включают взаимодействие с белками МНС. В результате селекции DP-тимоциты либо погибают (по прошествии примерно трех дней), либо становятся зрелыми лимфоцитами. При позитивной селекции происходит сшивание поверхностных αβ-гетеродимеров TCR, что предотвращает клеточную смерть тимоцитов. Если тимоциты выживают при последующей негативной селекции, они дают начало отдельным классам Т-лимфоцитов — $CD4^+CD8^-$ и $CD4^-CD8^+$.

Рецептор Т-клетки связан с комплексом белков **CD3**, задействованным в передаче сигнала с поверхнос-



РИС. 23.31 Развитие Т-клетки включает несколько последовательных этапов

ти внутрь клетки (в случае активации рецептора антигеном). Наши текущие представления о составе рецепторного комплекса Т-клетки иллюстрирует РИС. 23.32. Важно, что взаимодействие вариабельных участков TCR с антигеном побуждает ζ-субъединицы комплекса CD3 активировать ответ Т-клетки. Активация CD3 обеспечивает механизм, с помощью которого αβ или γδ TCR сообщает о том, что им распознан антиген. Это сравнимо со строением рецептора В-клетки, в которой иммуноглобулин взаимодействует с сигнальными цепями Igαβ (см. рис. 23.26).

К сожалению, ключевой вопрос, касающийся Т-клеток, пока не решен. Клеточно-опосредованный иммунитет требует друх процессов распознавания. Распознавание внешнего антигена предполагает способность реагировать на неизвестные структуры. Распознавание белка МНС, разумеется, ограничено тем, что закодировано в геноме, но существует множество различных белков МНС. Таким образом, обе реакции распознавания сталкиваются с существенным разнообразием. Хелперные и киллерные Т-клетки нуждаются в различных классах белков МНС, но для сборки рецепторов они используют один и тот же набор α- и β-сегментов гена. Даже если предположить внесение дополнительных вариаций в процессе рекомбинации TCR, остается неясным, каким образом формируется достаточное количество версий TCR, соответствующее всем этим требованиям.

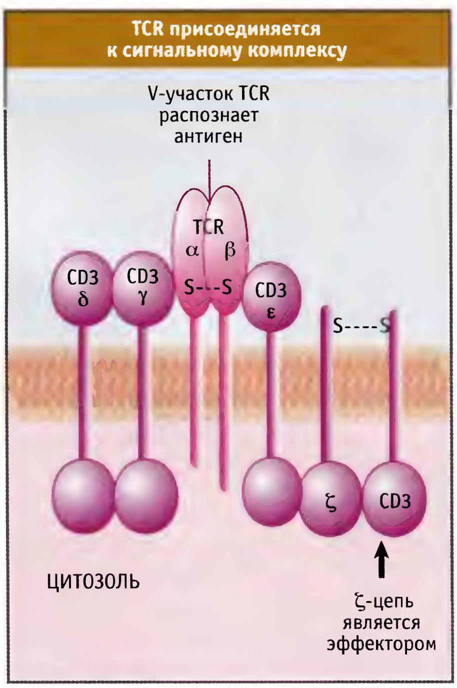


РИС. 23.32 Две цепи рецептора Т-клетки ассоциируют с полипептидами комплекса CD3. Вариабельные участки TCR экспонируются на поверхность клетки. Цитоплазменные домены ζ-цепей CD3 обеспечивают эффекторную функцию

23.20 Локус главного комплекса гистосовместимости кодирует множество генов иммунной системы

Ключевые понятия:

- Локус МНС кодирует белки класса I и II, а также другие белки иммунной системы.
- Белки класса I — антигены трансплантации, отвечающие за распознавание «своих» и «чужих» тканей.
- Белки МНС класса I активны в виде гетеродимера с β2-микроглобулином.
- Белки класса II работают при взаимодействии между Т-клетками.
- Белок МНС класса II представляет собой гетеродимер из α- и β-цепей.

Локус главного комплекса гистосовместимости (Major Histocompatibility Locus, MHC) занимает небольшой сегмент одной из хромосом как у мыши (где он носит название **локус H2**), так и у человека (локус **HLA**). Внутри этого сегмента расположено множество генов, кодирующих функции, относящиеся к иммунному ответу. В индивидуальных локусах генов, продукты которых были идентифицированы, в популяции было обнаружено множество аллелей. Этот локус описывают как высоко*полиморфный*, т. е. индивидуальные геномы обычно отличаются друг от друга по данному локусу. На этом участке также имеются гены, кодирующие функции, не связанные с работой иммунной системы.

По своим иммунологическим свойствам антигены гистосовместимости делят на три типа. Кроме того, некоторые белки, обнаруживаемые на поверхности лимфоцитов и макрофагов, имеют схожую структуру и играют важную роль в работе клеток иммунной системы:

Белки МНС класса I являются **трансплантационными антигенами**. Они присутствуют во всех клетках млекопитающих. Как следует из названия, они отвечают за отторжение чужеродной ткани, которая идентифицируется по уникальному набору трансплантационных антигенов. В иммунной системе их присутствие на целевых клетках необходимо для клеточно-опосредованного ответа. Типы белков класса I определяются серологически (на основании их антигенных свойств). Мышиные гены класса I кодируют белки H2-K и H2-D/L. Каждая линия мышей имеет один из возможных алелей для каждой из этих функций. У человека к функциям класса I относятся классические трансплантационные антигены HLA-A, B и C.

Белки МНС класса II присутствуют на поверхности как В- и Т-лимфоцитов, так и макрофагов. Эти белки включены в процессы межклеточных коммуникаций, необходимых для проведения иммунного ответа; в частности, они необходимы для функций Т-хелперов. У мышей класс II подразделяются на I-A и I-E. У человека участок HLA класса II (называемый также HLA-D) делят на четыре подучастка — DR, DQ, DZ/DO и DP.



23.33 Длина участка МНС – более 2 Мпн. Белки МНС классов I и II закодированы двумя различными участками. Разделяющий их сегмент определяется как класс III. Расширенные участки описывают синтеничные сегменты на концах кластера. Основное различие мыши и человека – наличие генов H2 класса I в расширенном участке слева. У мыши локус находится на хромосоме 17, а у человека – на хромосоме 6

Белки **комплемента** закодированы локусом, известным как S-участок (от «serum» — сыворотка; «S» указывает на то, что эти белки входят в состав сыворотки). Их назначение — взаимодействие с комплексом антиген-антитело, вызывая лизис клеток по классическому варианту гуморального ответа.

Белки локусов *Qa* и *Tla* обнаружены на гемопоэтических клетках мышей. Они известны как гены дифференцировки (поскольку каждый из них обнаруживается только на определенном подмножестве клеток крови). Структурно они напоминают белки H2 класса I, и, подобно последним, данные белки полиморфны.

Теперь мы можем установить соответствие между типами белков и организацией генов, их кодирующих. Изначально участок МНС был найден генетиками у мышей, у которых классический участок H2 занимает 0,3 единицы генетической карты. Вместе с примыкающим участком, где также обнаружены мутации, влияющие на иммунные функции, это соответствует участку ДНК длиной около 2000 тпн. Участок МНС был полностью секвенирован у нескольких млекопитающих, птиц и рыб. Сравнивая эти последовательности, можно увидеть, что их строение в целом сохраняется.

Строение генов человека и мыши показано на РИС. 23.33. Участки генома, где расположены гены классов I и II, отмечают исходные границы локуса (в направлении от теломеры к центромере — на рисунке справа налево). Гены участка, разделяющего класс I и класс II (называемого классом III), кодируют различные функции. Границы локуса различаются для разных видов. Участок, лежащий за генами класса I в сторону теломеры, носит название расширенного класса I. Аналогично, участок за пределами класса II в сторону центромеры, называют расширенным классом II. Основное отличие мыши от человека состоит в том, что у мыши расширенный класс II содержит гены класса I (H2-K).

У млекопитающих участок МНС содержит несколько сотен генов, но функции МНС могут обеспечиваться и меньшим числом генов. Например, у кур участок МНС длиной 92 тпн включает всего девять генов.

Сравнивая различные семейства, мы обнаруживаем различия в числе генов, ответственных за те или иные функции. Локус МНС обнаруживает значительные индивидуальные вариации, в результате которых число генов у разных особей может отличаться. Как правило, в геноме мыши меньше активных H2, чем у человека. Гены класса II встречаются только у млекопитающих (за исключением одной подгруппы), и, как правило, у птиц и рыб в данном локусе находятся различные гены. У человека имеется около 8 функциональных генов класса I, а у мыши — около 30. Участок класса I также включает много других генов. Участки класса III у человека и мыши очень похожи.

Все белки МНС являются димерами, расположенными на плазматической мембране, причем большая часть белка выступает с внешней стороны клетки. Структура белков класса I и II похожа, хотя они и состоят из разных компонентов (как показывает РИС. 23.34).

Антигены класса II состоят из двух цепей, α и β, сочетание которых создает общую структуру, в которой присутствуют два внеклеточных домена.

Все белки МНС класса I состоят из димера между самой цепью класса I и β2-микроглобулином. Цепь класса I — это 45 кДа трансмембранный компонент, состоящий из трех **внешних доменов** (длиной около 90 аминокислот каждый; один из них взаимодействует с β2-микроглобулином), **трансмембранного участка** длиной около 40 остатков и короткого **цитоплазматического домена** длиной около 30 остатков, находящегося внутри клетки.

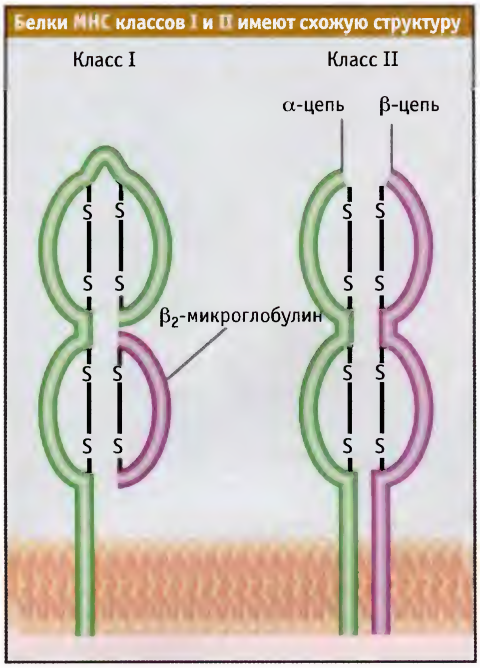


РИС. 23.34 **Антигены гистосовместимости классов I и II имеют схожую структуру. Антигены класса I состоят из единственного полипептида (α) с тремя внешними доменами (α1, α2, α3), взаимодействующего с микроглобулином β2 (β2 m). Антигены класса II состоят из двух полипептидов (α и β), каждый из который состоит из двух доменов (α1, α2 и β1, β2) с аналогичной общей структурой**

β2-микроглобулин — это 12-кДа секреторный белок, необходимый для транспортировки цепи класса I к поверхности клетки. У мышей, у которых нет гена β2-микроглобулина, на поверхности клетки отсутствует антиген МНС класса I.

Строение генов класса I, показанное на РИС. 23.35, совпадает с белковой структурой. Первый экзон кодирует сигнальную последовательность (отщепляемую от белка при прохождении через мембрану). Следующие три экзона кодируют каждый из внешних доменов. Пятый экзон кодирует трансмембранный домен. Последние три относительно небольших экзона в сумме кодируют цитоплазматический домен. Единственое отличие трансплантационных антигенов у человека (на уровне гена) состоит в том, что цитоплазматический домен закодирован лишь двумя экзонами.

Экзон, кодирующий третий внешний домен генов класса I, высококонсервативен (по сравнению с остальными экзонами). Консервативный домен предположительно представляет собой участок, взаимодействующий с β2-микроглобулином, что объясняет необходимость стабильности структуры. Этот домен также имеет сходство с константными участками доменов иммуноглобулинов.

Что является причиной высокой степени полиморфизма этих генов? Разнообразие последовательностей разных аллелей проявляется в основном в первом и втором внешних доменах, иногда принимая форму кластера с заменой оснований на небольшом участке. Один из механизмов, участвующих в их генерации, — это конверсия генов между генами класса I. Присутствуют как псевдогены, так и рабочие гены.

Ген β2-микроглобулина находится на отдельной хромосоме. В его состав входят четыре экзона: первый кодирует сигнальную последовательность, второй — основную часть белка (аминокислоты 3–95), третий — последние четыре аминокислоты и часть нетранслируемого окончания, четвертый — оставшуюся часть окончания.

Длина генов β2-микроглобулина и V-гена иммуноглобулина примерно одинакова; определенное сходство состава аминокислот и некоторое (небольшое) сходство нуклеотидной последовательности наблюдается между β2-микроглобулином и константными доменами Ig или третьими внешними доменами гена типа I. Все рассмотренные нами в этой главе группы генов, возможно, имеют общего предка, кодировавшего исходный домен.

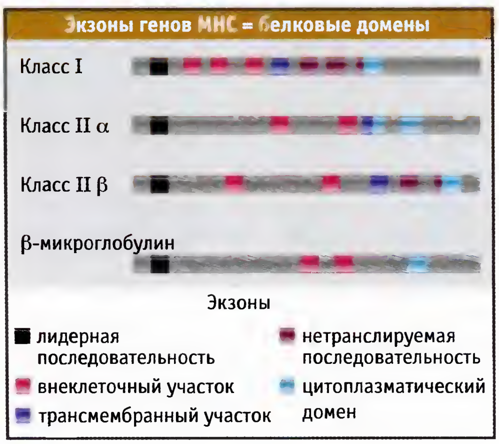


РИС. 23.35 **Каждый класс генов МНС имеет характерное строение, в котором экзоны представляют собой отдельные белковые домены**

23.21 Врожденный иммунитет в своей работе использует консервативные сигнальные пути

Ключевые понятия:

- Врожденный иммунитет активируется рецепторами (PAMP), распознающими определенные высококонсервативные мотивы у бактерий и у других инфекционных агентов.
- Toll-подобные рецепторы широко используются для активации этого пути.
- Система врожденного иммунитета высококонсервативна у беспозвоночных и позвоночных, аналогичная система обнаружена и у растений.

Разнообразие PAMP		
Организм	Патоген	Расположение
Все бактерии	Формил-метионин	Большинство белков
Большинство бактерий	Пептидогликан	Клеточная стенка
Грам-отрицательные бактерии	Липополисахарид	Клеточная стенка
Дрожжи	Зимозан	Клеточная стенка

РИС. 23.36 **Патоген-ассоциированные молекулярные последовательности (PAMP) – соединения, часто встречающиеся у широкого спектра видов бактерий и обнаруживаемые при инфицировании**

Иммунный ответ, описываемый в этой главе, состоит из набора реакций, отвечающих на проникновение патогена выбором лимфоцитов (В- или Т-клеток), рецепторы которых (антитела или TCR) высокоаффинны к патогену. Базой этого процесса является выработка очень большого числа рецепторов с целью повышения вероятности распознавания чужеродной молекулы. Рецепторы, распознающие собственные белки организма, выбраковываются на ранней стадии процесса. Активация рецепторов на В-клетках включает пути гуморального ответа, а активация рецепторов на Т-клетках включает пути клеточно-опосредованного ответа. Общая реакция на антиген путем отбора рецепторов на лимфоцитах носит название **адаптивного**, или **приобретенного иммунитета**. Ответ обычно развивается в течение нескольких дней после начальной активации В-клеток или Т-клеток, распознающих чужеродную последовательность. Организм сохраняет память об ответе, что позволяет ему быстрее реагировать на повторное появление того же патогена. Механизм приобретенного иммунитета является общим для всех позвоночных, хотя и существуют отличия в деталях.

Иной тип иммунного ответа развивается за меньшее время и присутствует у большего числа видов животных (включая и не имеющих приобретенного иммунитета). **Врожденный иммунитет** основывается на распознавании определенных, заранее заданных последовательностей, принадлежащих внешним патогенам. Эти последовательности являются фрагментами, которые постоянно присутствуют в патогенах, поскольку играют существенную роль в их жизнедеятельности, но не встречаются у высших эукариот. Как правило, фрагмент распознается рецептором, отвечающим за включение врожденного иммунного ответа при инфицировании. Рецепторы, включающие врожденный иммунный ответ, обнаружены на нейтрофилах и макрофагах, и с их помощью патоген поглощается и уничтожается. Ответ развивается быстро, поскольку набор рецепторов уже присутствует на клетках и нет необходимости в его амплификации путем отбора. Врожденный иммунитет широко распространен и имеется у многих организмов, от мухи до человека. Если врожденный иммунитет способен эффективно справиться с заражением, механизм приобретенного иммунитета не задействуется. Два вида ответа могут активировать одни и те же последовательности, поэтому клетки, активированные врожденным иммунитетом, могут впоследствии участвовать в адаптивном ответе.

Фрагменты, включающие врожденный ответ, иногда называют патоген-ассоциированными молекулярными последовательностями (pathogen-associated molecular patterns, **PAMP**). РИС. 23.36 показывает, что они встречаются у различных организмов. Формил-метионин является первой аминокислотой у большинства белков бактерий, но не встречается у эукариот. Пептидогликан имеется только у бактерий (в клеточной стенке). **Липополисахариды** (другое название — **эндотоксины**) входят в состав внешней мембраны большинства грам-отрицательных бактерий; они являются причиной синдрома септического шока.

Важнейшим открытием в области природы врожденного иммунитета было обнаружение участия в нем **рецепторов TLR**. Рецептор Toll, родственный рецептору IL1 у млекопитающих, у *Drosophila* активирует последовательность, управляющую дорсо-вентральным

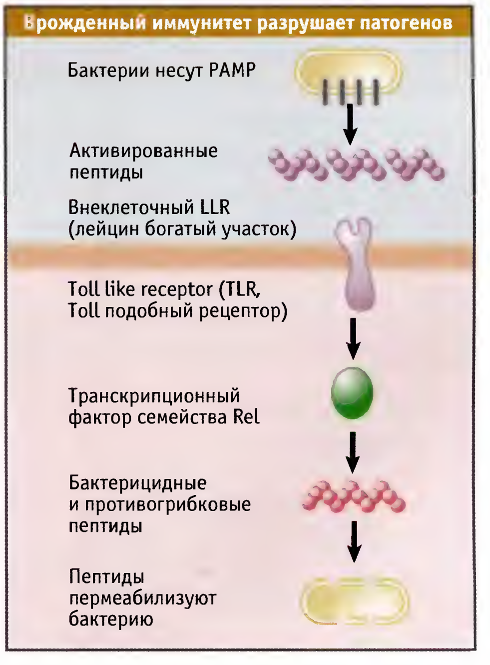
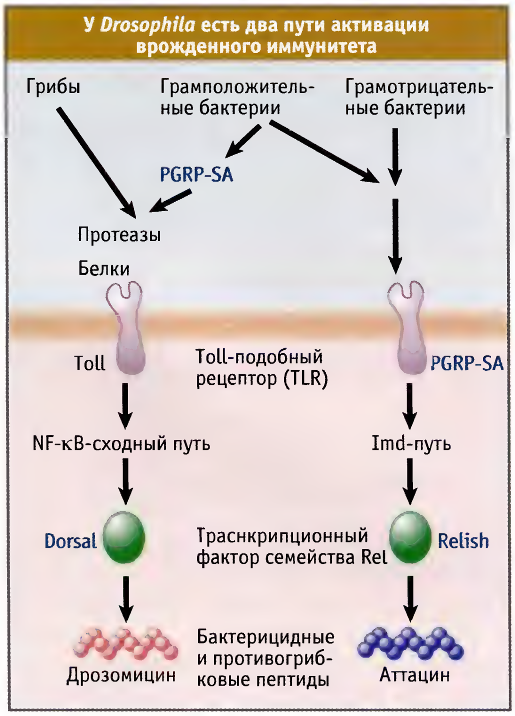


РИС. 23.37 **Врожденный иммунитет активизируется PAMP. У мух они вызывают продукцию пептидов, активирующих рецепторы TLR. Рецепторы TLR включают путь, активирующий транскрипционный фактор семейства Rel. В числе целевых генов для этого фактора – бактерицидные и протиогрибковые пептиды. Данные пептиды пермеабилизуют мембрану патогенного организма**

развитием. Это ведет к активации фактора, входящего в группу Rel (родственную фактору NF-κB у млекопитающих). Путь врожденного иммунитета сходен с последовательностью Toll, с аналогичными компонентами. По существу, одним из первых сведений о природе врожденного иммунитета у мух было открытие фактора Dif (дорсо-связанного фактора иммунитета), который активируется в одном из путей.

У мух отсутствует система приобретенного иммунитета, но имеется сопротивляемость микробному заражению, поскольку их системы врожденного иммунитета вырабатывают мощные противомикробные пептиды. У дрозофилы обнаружены семь различных пептидов, синтезируемых жировым телом (орган, эквивалентный печени). Два из них противогрибковые, а пять действуют преимущественно на бактерий. Они уничтожают чужеродный организм, пермеабилизуя его клеточную мембрану. Все эти пептиды кодируются генами, промоторы которых отвечают на факторы транскрипции семейства Rel. Схемы путей врожденного иммунного ответа у дрозофилы показаны на РИС. 23.37.

Существует два пути врожденного иммунного ответа у *Drosophila*; один реагирует в основном на грибы, другой — на грамотрицательные бактерии. На грам-положительные бактерии могут реагировать оба пути. Их схема показана на РИС. 23.38. Грибы и грамположительные бактерии активируют протеолитический каскад, создающий пептиды, которые



3.38 Один из путей активации врожденного иммунитета у *Drosophila* очень схож с путем активации NF-κB у млекопитающих; другие пути используют компоненты, аналогичные участвующим в апоптозе

активируют TLR. Это NF-κB-подобная последовательность. Рецептор dToll активирует транскрипционный фактор Dif (родственный NF-κB), что в итоге приводит к активации противогрибкового пептида дрозомицина. Грам-отрицательные бактерии задействуют иной путь, воздействующий на другой рецептор, активирующий транскрипционный фактор Relish, влияющий на выработку бактерицидного пептида аттацина. Этот путь носит название Imd (по одному из своих компонентов — белку, имеющему «домен смерти», схожему с участвующим в апоптозе).

Основные агенты, реагирующие на бактерий, — это белки, называемые PGPR (из-за их высокой аффинности к бактериальным пептидогликанам). Существуют два вида этих белков. PGPR-SA — это короткие внеклеточные белки. Их механизм действия, вероятно, основан на активации протеаз, включающих Toll-путь. PGPR-LC — это трансмембранные белки, имеющие внеклеточный домен PGPR. Их роль пока не установлена.

Врожденный иммунитет — высококонсервативная система. Мыши, устойчивые к септическому шоку при введении LPS, несут мутации в TLR4. Гомолог Toll-рецептора у человека может активировать некоторые гены иммунного ответа. Это говорит о том, что данный путь врожденного иммунитета также может функционировать у человека. Сигнальный путь от TLR во всех случаях похож и обычно приводит к активации транскрипционого фактора NF-κB. Неизвестно, консервативен ли сигнальный путь до TLR, в частности действуют ли PAMP путем образования лигандов, которые в свою очередь активируют TLR, или они непосредственно взаимодействуют с TLR. Сигнальный путь до TLR различен у млекопитающих и мух, поскольку патогены напрямую активируют TLR млекопитающих. В случае LPS патоген связывается с поверхностным белком CD14; это позволяет CD14 активировать TLR4, запуская путь активации врожденного иммунитета. В геноме человека около 20 рецепторов класса TLR, что дает некоторое представление о том, сколько патогенов могут активировать врожденный иммунитет.

У растений имеются многочисленные защитные механизмы, среди которых есть сигнальные пути, аналогичные врожденному иммунитету животных. Действует тот же принцип: PAMP — это фрагменты, которые идентифицируют инфекционный агент как патоген. Белки, отвечающие на патогены, закодированы классом генов, носящих название «гены устойчивости к заболеванию» (disease resistance genes). Многие из этих генов кодируют рецепторы, имеющие общие свойства с рецепторами TLR-класса у животных: внеклеточный домен несет мотив, называемый **LRR** (от *англ.* leucine-rich region — **участок, богатый лейцином**). Механизм ответа отличается от имеющегося у клеток животных и направлен на активацию каскада MAPK (mitogen-activated protein kinase, митоген-активируемая протеинкиназа). Разные патогены активируют один и тот же каскад, свидетельствуя о том, что разнообразные взаимодействия патоген-рецептор сходятся в одной точке.

23.22 Резюме

Иммуноглобулины и рецепторы Т-клеток — это белки, выполняющие схожие функции у В- или Т-клеток иммунной системы. Белки Ig или TCR образуются путем реорганизации ДНК в каждом конкретном лимфоците; встреча с антигеном и распознавание его с помощью Ig или TCR приводит к клональной экспансии и образованию множества клеток с той же специфичностью, что и исходная клетка. В раннем развитии иммунной системы происходит множество различных реорганизаций, что создает широкий репертуар клеток с разнообразными специфичностями.

Каждый белок-иммуноглобулин — это тетрамер, содержащий две идентичные легкие цепи и две идентичные тяжелые цепи. TCR — это димер, содержащий две разные цепи. Каждая полипептидная цепь — результат экспрессии гена, создаваемого путем соединения одного из многих V-сегментов с D- и J-сегментами и с одним из нескольких С-сегментов. L-цепи (κ или λ) Ig имеют общую структуру V-J-C, Н-цепи Ig имеют структуру V-D-J-C, TCR α или γ имеют составные части, аналогичные L-цепи Ig, а TCR δ и β — аналогичные Н-цепи Ig.

Каждый тип цепей закодирован большим кластером V-генов, отделенным от кластеров D-, J- и С-сегментов. Число сегментов каждого типа и их организация у разных цепей отличаются, но принцип и механизм рекомбинации, по-видимому, одинаковы. В каждом акте рекомбинации участвуют те же самые консенсусные последовательности гептамера и нонамера; и в реакции всегда соединяются консенсусная последовательность со спейсером в 23 пн и со спейсером в 12 пн. Реакция расщепления катализируется белками RAG1 и RAG2, а реакция соединения катализируется участниками одного и того же пути NHEJ, репарирующими в клетке двухцепочечные разрывы. Механизм действия белков RAG сходен с действием сайт-специфичной рекомбинации, катализируемой резольвазами.

Существенное разнообразие антител достигается при соединении различных V-, D- и J-сегментов с С-сегментом; при этом дополнительные вариации вносятся в форме изменений в местах соединения сегментов во время процесса рекомбинации. Дополнительные изменения в гены иммуноглобулинов вносятся путем соматического мутагенеза, для чего необходима активность цитидин-дезаминазы и урацил-гликозилазы. Мутации, индуцируемые цитидиндезаминазой, вероятно, приводят к удалению урацила урацил-гликозилазой, что сопровождается индукцией мутаций в сайте с отсутствующим азотистым основанием.

Аллельное исключение гарантирует, что каждый конкретный лимфоцит синтезирует только один Ig или TCR. Продуктивная реорганизация ингибирует дальнейшую реорганизацию. Первая продуктивная реорганизация определяет использование конкретного V-участка, но в В-клетках происходит переключение использования C_H-генов с исходной μ-цепи на одну из Н-цепей, кодируемую ниже цепи μ. В этом процессе используется другой тип рекомбинации, при котором последовательности между VDJ-участком и новым C_H-геном делетируются. Может происходить более чем одно переключение использования C_H-генов. При переключении классов нужна та же цитидин-дезаминаза, которая необходима и для соматического мутагенеза, однако ее роль в данном процессе неизвестна. На ранней стадии продукции Ig переключения происходят между синтезом мембраносвязанной формы белка к синтезу секретируемой формы. Эти переключения сопровождаются альтернативным сплайсингом транскрипта.

Врожденный иммунитет представляет собой тип ответа, запускаемый рецепторами, специфичность которых к некоторым часто встречающимся мотивам у бактерий и других инфекционных агентов предопределена. Рецепторы, запускающие данный тип ответа, — это обычно члены Toll-подобного семейства, и их путь напоминает путь ответа на сигналы Toll-рецепторов при эмбриональном развитии. Сигнальный путь активирует факторы транскрипции, приводящие к экспрессии генов, продукты которых инактивируют инфекционных агентов (обычно путем пермеабилизации их мембраны).

Литература

23.3 Гены иммуноглобулинов компонуются из составных частей в лимфоцитах

Обзор

Alt, F. W., Blackwell, T. K., and Yancopoulos, G. D. (1987). Development of the primary antibody repertoire. *Science* 238, 1079–1087.

Blackwell, T. K. and Alt, F. W. (1989). Mechanism and developmental program of immunoglobulin gene rearrangement in mammals. *Annu. Rev. Genet.* 23, 605–636.

Hood, L., Kronenberg, M., and Hunkapiller, T. (1985). T cell antigen receptors and the immunoglobulin supergene family. *Cell* 40, 225–229.

Tonegawa, S. (1983). Somatic generation of antibody diversity. *Nature* 302, 575–581.

Yancopoulos, G. D. and Alt, F. W. (1986). Regulation of the assembly and expression of variable-region genes. *Annu. Rev. Immunol.* 4, 339–368.

Исследование

Hozumi, N. and Tonegawa, S. (1976). Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proc. Natl. Acad. Sci. USA* 73, 3628–3632.

23.4 Легкие цепи образуются путем единственного акта рекомбинации

Исследование

Max, E. E., Seidman, J. G., and Leder, P. (1979). Sequences of five potential recombination sites encoded close to an immunoglobulin k constant region gene. *Proc. Natl. Acad. Sci. USA* 76, 3450–3454.

23.7 При рекомбинации генов иммунной системы используются два типа консенсусных последовательностей

Исследование

Lewis, S., Gifford, A., and Baltimore, D. (1985). DNA elements are asymmetrically joined during the site-specific recombination of kappa immunoglobulin genes. *Science* 228, 677–685.

23.9 Аллельное исключение инициируется продуктивной реорганизацией генома

Обзор

Storb, U. (1987). Transgenic mice with immunoglobulin genes. *Annu. Rev. Immunol.* 5, 151–174.

23.10 Белки RAG катализируют разрыв и воссоединение ДНК

Обзор

Gellert, M. (1992). Molecular analysis of VDJ recombination *Annu. Rev. Genet.* 26, 425–446.

Jeggo, P. A. (1998). DNA breakage and repair. *Adv. Genet.* 38, 185–218.

Schatz, D. G., Oettinger, M. A., and Schlissel, M. S. (1992). VDJ recombination: molecular biology and regulation. *Annu. Rev. Immunol.* 10, 359–383.

Исследование

Agrawal, A., Eastman, Q. M., and Schatz, D. G. (1998). Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. *Nature* 394, 744–751.

Hiom, K., Melek, M., and Gellert, M. (1998). DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. *Cell* 94, 463–470.

Ma, Y., Pannicke, U., Schwarz, K., and Lieber, M. R. (2002). Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell* 108, 781–794.

Melek, M. and Gellert, M. (2000). RAG1/2-mediated resolution of transposition intermediates: two pathways and possible consequences. *Cell* 101, 625–633.

Qiu, J. X., Kale, S. B., Yarnell Schultz, H., and Roth, D. B. (2001). Separation-of-function mutants reveal critical roles for RAG2 in both the cleavage and joining steps of V(D)J recombination. *Mol. Cell* 7, 77–87.

Roth, D. B., Menetski, J. P., Nakajima, P. B., Bosma, M. J., and Gellert, M. (1992). V D J recombination: broken DNA molecules with covalently sealed (hairpin) coding ends in SCID mouse thymocytes. *Cell* 70, 983–991.

Schatz, D. G. and Baltimore, D. (1988). Stable expression of immunoglobulin gene V(D)J recombinase activity by gene transfer into 3T3 fibroblasts. *Cell* 53, 107–115.

Schatz, D. G., Oettinger, M. A., and Baltimore, D. (1989). The V(D)J recombination activating gene, RAG-1. *Cell* 59, 1035–1048.

Tsai, C. L., Drejer, A. H., and Drejer, A. H. (2002). Evidence of a critical architectural function for the RAG proteins in end processing, protection, and joining in V(D)J recombination. *Genes Dev.* 16, 1934–1949.

Yarnell Schultz, H., Landree, M. A., Qiu, J. X., Kale, S. B., and Roth, D. B. (2001). Joining-deficient RAG1 mutants block V(D)J recombination *in vitro* and hairpin opening *in vitro*. *Mol. Cell* 7, 65–75.

23.13 Переключение происходит путем необычной реакции рекомбинации

Обзор

Honjo, T., Kinoshita, K., and Muramatsu, M. (2002). Molecular mechanism of class switch recombination: linkage with somatic hypermutation. *Annu. Rev. Immunol.* 20, 165–196.

Li, Z., Woo, C. J., Iglesias-Ussel, M. D., Ronai, D., and Scharff, M. D. (2004). The generation of antibody diversity through somatic hypermutation and class switch recombination. *Genes Dev.* 18, 1–11.

Исследование

Bransteitter, R., Pham, P., Scharff, M. D., and Goodman, M. F. (2003). Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of PHKse. *Proc. Natl. Acad. Sci. USA* 100, 4102–4107.

Gu, H., Zou, Y. R., and Rajewsky, K. (1993). Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. *Cell* 73, 1155–1164.

Iwasato, T., Shimizu, A., Honjo, T., and Yamagishi, H. (1990). Circular DNA is excised by immunoglobulin class switch recombination. *Cell* 62, 143–149.

Kinoshita, K., Tashiro, J., Tomita, S., Lee. C. G., and Honjo, T. (1998). Target specificity of immunoglobulin class switch recombination is not determined by nucleotide sequences of S regions. *Immunity* 9, 849–858.

Manis, J. P., Gu, Y., Lansford, R., Sonoda, E., Ferrini, R., Davidson, L., Rajewsky, K., and Alt, F. W. (1998). Ku70 is required for late B cell development and immunoglobulin heavy chain class switching. *J. Exp. Med.* 187, 2081–2089.

Matsuoka, M., Yoshida, K., Maeda, T., Usuda, S., and Sakano, H. (1990). Switch circular DNA formed in cytokine-treated mouse splenocytes: evidence for intramolecular DNA deletion in immunoglobulin class switching. *Cell* 62, 135–142.

Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., and Honjo, T. (2000). Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential PHK editing enzyme. *Cell* 102. 553–563.

Pham, P., Bransteitter, R., Petruska, J., and Goodman, M. F. (2003). Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature* 424, 103–107.

Revy, P., Muto, T., Levy, Y., Geissmann, F., Plebani, A., Sanal, O., Catalan, N., Forveille, M., Dufourcq-Labelouse, R., Gennery, A., Tezcan. I., Ersoy, F.. Kayserili, H., Ugazio, A. G., Brousse, N., Muramatsu, M., Notarangelo, L. D., Kinoshita, K., and Honjo, T. (2000). Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* 102, 565–575.

Rolink, A., Melchers, F., and Andersson, J. (1996). The SCID but not the RAG-2 gene product is required for S mu-S epsilon heavy chain class switching. *Immunity* 5, 319–330.

von Schwedler, U., Jack, H. M., and Wabl, M. (1990). Circular DNA is a product of the immunoglobulin class switch rearrangement. *Nature* 345, 452–456.

Xu, L., Gorham, B., Li, S. C., Bottaro, A., Alt, F. W., and Rothman, P. (1993). Replacement of germ-line epsilon promoter by gene targeting alters control of immunoglobulin heavy chain class switching. *Proc. Natl. Acad. Sci. USA* 90, 3705–3709.

23.14 Соматические мутации создают дополнительное многообразие антител у мышей и человека

Обзор

French, D. L., Laskov, R., and Scharff, M. D. (1989). The role of somatic hypermutation in the generation of antibody diversity. *Science* 244, 1152–1157.

Kocks, C. and Rajewsky, K. (1989). Stable expression and somatic hypermutation of antibody V regions in B-cell developmental pathways. *Annu. Rev. Immunol.* 7, 537–559.

Исследование

Kim, S., Davis, M., Sinn, E., Patten, P., and Hood, L. (1981). Antibody diversity: somatic hypermutation of rearranged VH genes. *Cell* 27, 573–581.

23.15 Цитидин-дезаминаза и урацил-гликозилаза вызывают соматический мутагенез

Обзор

Honjo, T., Kinoshita, K., and Muramatsu, M. (2002). Molecular mechanism of class switch recombination: linkage with somatic hypermutation. *Annu. Rev. Immunol.* 20, 165–196.

Kinoshita, K. and Honjo, T. (2001). Linking class-switch recombination with somatic hypermutation. *Nat. Rev. Mol. Cell Biol.* 2, 493–503.

Li, Z., Woo, C. J., Iglesias-Ussel, M. D., Ronai, D., and Scharff, M. D. (2004). The generation of antibody diversity through somatic hypermutation and class switch recombination. *Genes Dev.* 18, 1–11.

Исследование

Di Noia, J. and Neuberger, J. (2002). Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. *Nature* 419, 43–48.

Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., and Honjo, T. (2000). Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential PHK editing enzyme. *Cell* 102, 553–563.

Peters, A. and Storb, U. (1996). Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. *Immunity* 4, 57–65.

Revy, P., Muto, T., Levy, Y., Geissmann, F., Plebani, A., Sanal, O., Catalan, N., Forveille, M., Dufourcq-Labelouse, R., Gennery, A., Tezcan, I., Ersoy, F., Kayserili, H., Ugazio, A. G., Brousse, N., Muramatsu, M., Notarangelo, L. D., Kinoshita, K., and Honjo, T. (2000). Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* 102, 565–575.

23.16 Иммуноглобулины птиц образуются из псевдогенов

Исследование

Reynaud, C. A., Anquez, V., Grimal, H., and Weill, J. C. (1987). A hyperconversion mechanism generates the chicken light chain preimmune repertoire. *Cell* 48, 379–388.

Sale, J. E., Calandrini, D. M., Takata, M., Takeda, S., and Neuberger, M. S. (2001). Ablation of XRCC2/3 transforms immunoglobulin V gene conversion into somatic hypermutation. *Nature* 412, 921–926.

23.17 Быстрый вторичный ответ возможен благодаря В-клеткам памяти

Обзор

Rajewsky, K. (1996). Clonal selection and learning in the antibody system. *Nature* 381, 751–758.

23.18 Рецепторы Т-клеток схожи с иммуноглобулинами

Обзор

Davis, M. M. (1990). T-cell receptor gene diversity and selection. *Annu. Rev. Biochem.* 59, 475–496.

Kronenberg, M., Siu, G., Hood, L. E., and Shastri, N. (1986). The molecular genetics of the T-cell antigen receptor and T-cell antigen recognition. *Annu. Rev. Immunol.* 4, 529–591.

Marrack, P. and Kappler, J. (1987). The T-cell receptor. *Science* 238, 1073–1079.

Raulet, D. H. (1989). The structure, function, and molecular genetics of the gamma/delta T-cell receptor. *Annu. Rev. Immunol.* 7, 175–207.

23.19 Рецепторы Т-клеток действуют совместно с МНС

Обзор

Goldrath, A. W. and Bevan, M. J. (1999). Selecting and maintaining a diverse T cell repertoire. *Nature* 402, 255–262.

23.20 Локус главного комплекса гистосовместимости кодирует множество генов иммунной системы

Обзор

Flavell, R. A., Allen, H., Burkly, L. C., Sherman, D. H., Waneck, G. L., and Widera, G. (1986). Molecular biology of the H-2 histocompatibility complex. *Science* 233, 437–443.

Kumnovics, A., Takada, T., and Lindahl, K. F. (2003). Genomic organization of the mammalian MHC. *Annu. Rev. Immunol.* 21, 629–657.

Steinmetz, M. and Hood, L. (1983). Genes of the MHC complex in mouse and man. *Science* 222, 727–732.

Исследование

Kaufman, J. et al. (1999). The chicken B locus is a minimal essential major histocompatibility complex. *Nature* 401, 923–925.

23.21 Врожденный иммунитет в своей работе использует консервативные сигнальные пути

Обзор

Aderem, A. and Ulevitch, R. J. (2000). Toll-like receptors in the induction of the innate immune response. *Nature* 406, 782–787.

Dangl, J. L. and Jones, J. D. (2001). Plant pathogens and integrated defence responses to infection. *Nature* 411, 826–833.

Hoffmann, J. A., Kafatos, F. C., Janeway, C. A., and Ezekowitz, R. A. (1999). Phylogenetic perspectives in innate immunity. *Science* 284, 1313–1318.

Janeway, C. A. and Medzhitov, R. (2002). Innate immune recognition. *Annu. Rev. Immunol.* 20, 197–216.

Исследование

Asai, T., Tena, G., Plotnikova, J., Willmann, M. R., Chiu, W. L., Gomez-Gomez, L., Boller, T., Ausubel, F. M., and Sheen, J. (2002). MAP kinase signalling cascade in Arabidopsis innate immunity. *Nature* 415, 977–983.

Hoffmann, J. A. (2003). The immune response of *Drosophila*. *Nature* 426, 33–38.

Ip, Y. T., Reach, M., Engstrom, Y., Kadalayil, L., Cai, H., González-Crespo, S., Tatei, K., and Levine, M. (1993). Dif, a dorsal-related gene that mediates an immune response in *Drosophila*. *Cell* 75, 753–763.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J. M., and Hoffmann, J. A. (1996). The dorsoventral regulatory gene cassette spÃtzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. *Cell* 86, 973–983.

Medzhitov, R., Preston-Hurlburt, P., and Janeway, C. A. (1997). A human homologue of the *Drosophila* Toll protein signals activation of adaptive immunity. *Nature* 388, 394–397.

Poltorak, A., He, X., Smirnova, I., Liu, M. Y., Huffel, C. V., Du, X., Birdwell, D., Alejos, E., Silva, M., Galanos, C., Freudenberg, M., Ricciardi-Castagnoli, P., Layton, B., and Beutler, B. (1998). Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science* 282, 2085–2088.

Rutschmann, S., Jung, A. C., Hetru, C., Reichhart, J. M., Hoffmann, J. A., and Ferrandon, D. (2000). The Rel protein DIF mediates the antifungal but not the antibacterial host defense in *Drosophila*. *Immunity* 12, 569–580.

Williams, M. J., Rodriguez, A., Kimbrell, D. A., and Eldon, E. D. (1997). The 18-wheeler mutation reveals complex antibacterial gene regulation in *Drosophila* host defense. *EMBO J.* 16, 6120–6130.

Промоторы и энхансеры

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

24.1 Введение

24.2 Эукариотическая ДНК-полимераза состоит из многих субъединиц

- РНК-полимераза I синтезирует рРНК в ядрышке.
- РНК-полимераза II синтезирует мРНК в нуклеоплазме.
- РНК-полимераза III синтезирует малые РНК в нуклеоплазме.
- Все эукариотические РНК-полимеразы имеют ~12 субъединиц и представляют собой комплексы с массой >500 кДа.
- Некоторые субъединицы являются общими для всех трех РНК-полимераз.
- Самая крупная субъединица РНК-полимеразы II несет CTD (С-концевой домен), состоящий из множественных повторов гептамера.

24.3 Промоторные элементы определяют с помощью мутагенеза и футпринтинга

- Промоторы выявляют по их способности вызывать транскрипцию присоединенной к ним последовательности в соответствующей тест-системе *in vitro* и *in vivo*.

24.4 Промотор РНК-полимеразы I состоит из двух частей

- Промотор РНК-полимеразы I состоит из центральной (core) области и контрольного элемента, расположенного выше (UPE, upstream control element).
- Фактор UBF1 накручивает ДНК вокруг белковой структуры, чтобы сблизить центральную область промотора и UPE.
- SL1 включает в себя фактор TBP, который используется при инициации всеми тремя РНК-полимеразами.
- РНК-полимераза связывается с комплексом UBF1-SL1 в центральной области промотора.

24.5 РНК-полимераза III использует и вышележащие, и нижележащие промоторы

- РНК-полимераза III узнает два типа промоторов.
- Внутренние промоторы имеют короткие консенсусные последовательности, расположенные в транскрипционной единице, и вызывают инициацию на определенном расстоянии выше промотора.
- Вышележащие промоторы содержат три коротких консенсусных последовательности, связывающих транскрипционные факторы.

24.6 $TF_{III}B$ – обязательный фактор для промоторов РНК-полимеразы III

- $TF_{III}A$ и $TF_{III}C$ связываются с консенсусными последовательностями и способствуют связыванию $TF_{III}B$ со стартовой точкой.
- $TF_{III}B$ несет TBP в качестве одной из субъединиц и делает возможным связывание РНК-полимеразы.

24.7 Стартовая точка для РНК-полимеразы II

- Для инициации транскрипции РНК-полимеразе II необходимо присутствие главных транскрипционных факторов (называемых $TF_{II}X$).
- Промоторы РНК-полимеразы II имеют короткую консервативную последовательность Py_2CAPy_5 (инициатор InR) в стартовой точке.
- TATA-бокс — общий компонент промоторов для РНК-полимеразы II, состоящий из А-Т-богатого октамера, расположенного на 25 пн выше стартовой точки.
- DPE — общий компонент для промоторов РНК-полимеразы II, не несущих TATA-бокса.
- Главный элемент промотора для РНК-полимеразы II включает в себя InR и либо TATA-бокс, либо DPE.

24.8 TBP — универсальный фактор

- TBP — компонент направляющего фактора, необходимого для связывания каждого типа РНК-полимеразы с ее промотором.
- Этим фактором для РНК-полимеразы II является $TF_{II}D$, состоящий из TBP и 11 TAF, общей массой ~800 кДа.

24.9 TBP связывает ДНК необычным способом

- TBP связывается с TATA-боксом в малой бороздке ДНК.
- Он полуокружает ДНК и изгибает ее примерно на 80°.
- Некоторые из TAF напоминают гистоны и могут образовывать структуры, напоминающие гистоновый октамер (нуклеосому).

24.10 Основной аппарат собирается на промоторе

- Первый шаг в инициации – связывание $TF_{II}D$ с TATA-боксом.
- Остальные факторы транскрипции связываются с комплексом в определенном порядке, увеличивая длину закрытого участка на ДНК.
- Когда РНК-полимераза II связывается с таким комплексом, она инициирует транскрипцию.

24.11 Инициация сопровождается освобождением промотора

- Для плавления ДНК и движения полимеразы необходимы $TF_{II}E$ и $TF_{II}H$.
- Для начала элонгации может потребоваться фосфорилирование CTD.
- Дальнейшее фосфорилирование CTD необходимо некоторым промоторам для завершения неудачной инициации.
- CTD может координировать процессинг РНК с транскрипцией.

24.12 Связь между транскрипцией и репарацией

- Транскрибируемые гены в случае повреждения ДНК быстро репарируются.
- $TF_{II}H$ обеспечивает связь с комплексом ферментов репарации.
- Мутации компонента XPD в $TF_{II}H$ у человека вызывают три типа заболеваний.

24.13 Активаторы связываются с короткими элементами последовательности

- Короткие консервативные элементы последовательности распределены на участках, предшествующих стартовой точке.
- Вышележащие элементы увеличивают частоту инициации.
- Факторы, которые связываются с ними для стимуляции транскрипции, называют активаторами.

24.14 Конструкция промотора может видоизменяться, но контекст имеет значение

- Ни один из вышележащих элементов не является необходимым для функционирования промоторов, тем не менее хотя бы один элемент обязательно должен присутствовать для эффективной инициации.
- Некоторые элементы распознаются множеством факторов, и фактор, который используется каждым конкретным промотором, можно определить в контексте других, связывающихся с промотором факторов.

24.15 Энхансеры несут двунаправленные элементы, способствующие инициации

- Энхансер активирует ближайший к нему промотор, расположенный на любом расстоянии выше или ниже него.
- UAS (upstream activator sequence, вышележащая активаторная последовательность) у дрожжей похожа на энхансер, но работает только выше промотора.
- В энхансерах и промоторах обнаруживаются схожие элементы последовательности.
- Энхансеры образуют комплексы с активаторами, которые прямым или косвенным образом взаимодействуют с промоторами.

24.16 Энхансеры содержат те же элементы, что и промоторы

- Энхансеры содержат те же самые короткие элементы последовательности, которые обнаружены в промоторах.
- Плотность элементов последовательности в энхансерах выше, чем в промоторах.

24.17 Функция энхансеров состоит в увеличении концентрации активаторов возле промотора

- Обычно энхансеры функционируют только в *цис*-положении относительно промотора-мишени.
- Можно заставить энхансеры работать и в *транс*-положении, соединив ДНК, несущую целевой промотор, с ДНК, несущей энхансер, через белковый мостик, или если эти две кольцевые молекулы будут сцеплены друг с другом.
- Суть в том, что энхансер будет функционировать в любом случае, если он находится недалеко от промотора.

24.18 Экспрессия генов связана с деметилированием

- Деметилирование 5'-конца гена необходимо для транскрипции.

24.19 CpG-островки являются мишенями для регуляции

- Неметилированные CpG-островки окружают промоторы конститутивно экспрессирующихся генов.
- CpG-островки обнаруживаются также в промоторах некоторых тканеспецифичных генов.
- В геноме человека примерно 29 000 CpG-островков.
- Метилирование CpG-островка не позволяет промотору внутри него активироваться.
- Репрессия вызывается белками, связывающимися с CpG-динуклеотидами.

24.20 Резюме

24.1 Введение

Для инициации транскрипции необходимы фермент — РНК-полимераза — и транскрипционные факторы. Под факторами транскрипции подразумеваются любые белки, необходимые для запуска транскрипции, но не являющиеся частью РНК-полимеразы. Многие факторы транскрипции действуют путем распознавания *цис*-действующих участков ДНК. Однако связывание с ДНК является не единственным механизмом действия транскрипционного фактора. Фактор может распознавать другой фактор или РНК-полимеразу, а также входить в состав инициаторного комплекса только при наличии ряда других белков. Ключевой признак участия в транскрипционном аппарате — функциональный: наличие белка должно быть обязательным условием прохождения транскрипции у данного промотора или группы промоторов.

Существенным различием в транскрипции эукариотической и прокариотической мРНК яляется то, что в инициации эукариотического промотора участвует большое число факторов, связывающихся с различными *цис*-действующими элементами. Промотор определяется как фрагмент, включающий все связывающие сайты, способные поддерживать транскрипцию с нормальной эффективностью и надлежащей регуляцией. Таким образом, основным фактором, определяющим эукариотический промотор, является расположение связывающих сайтов для факторов транскрипции. Сама по себе РНК-полимераза связывается в сайте начала транскрипции, но не контактиру-

ет напрямую с вышележащим участком промотора. В противоположность эукариотам, бактериальные промоторы, описанные в гл. 11 «Транскрипция», обычно определяются по расположению связывающего сайта РНК-полимеразы в непосредственной близости к точке начала транскрипции.

Транскрипцию в эукариотических клетках делят на три класса. Каждый класс транскрибируется своим видом РНК-полимеразы:

- РНК-полимераза I транскрибирует рРНК,
- РНК-полимераза II транскрибирует мРНК,
- РНК-полимераза III транскрибирует тРНК и другие небольшие РНК.

Транскрипционные факторы необходимы для инициации транскрипции, но впоследствии не требуются. За распознавание промотора отвечают дополнительные *факторы*, а не сами РНК-полимеразы. В этом отличие от бактериальной РНК-полимеразы, где, напротив, сам фермент распознает промоторную последовательность. Для всех эукариотических РНК-полимераз факторы создают такую структуру на промоторе, чтобы она могла быть распознана ферментом. Для РНК-полимераз I и III эти факторы устроены относительно просто, а для РНК-полимеразы II они образуют большую группу, и в комплексе известны как **основные факторы**. Основные факторы соединяются с РНК-полимеразой II для образования комплекса, окружающего стартовую точку транскрипции, и определяют сайт инициации. Основные факторы вместе с РНК-полимеразой образуют **основной транскрипционный аппарат**.

Промоторы для РНК-полимераз I и II чаще всего расположены выше стартовой точки начала транскрипции, однако некоторые промоторы для РНК-полимеразы III лежат ниже стартовой точки. Каждый промотор содержит характерный набор коротких консервативных последовательностей, распознаваемых соответствующим классом факторов. РНК-полимеразы I и III способны распознавать довольно ограниченный набор промоторов и зависят от небольшого числа вспомогательных факторов.

Промоторы, распознаваемые РНК-полимеразой II, обладают большим разнообразием последовательностей и имеют модульную организацию. Короткие элементы последовательностей, распознаваемые факторами транскрипции, находятся выше стартовой точки. Эти *цис*-действующие сайты обычно находятся на участке >200 пн. Некоторые из этих элементов и факторов, их распознающих, являются универсальными: они задействованы в работе различных промоторов и используются конститутивно. Но есть и специфичные, которые распознают отдельные классы генов и работа которых регулируется. Эти элементы встречаются в различных комбинациях в некоторых промоторах.

Все промоторы для РНК-полимеразы II имеют специальные элементы последовательности (неподалеку от точки начала транскрипции), связываемые основным аппаратом и определяющие сайт инициации. Более далекие от них последовательности определяют, экспрессируется данный промотор во всех типах клеток или регулируется специфически. Конститутивно

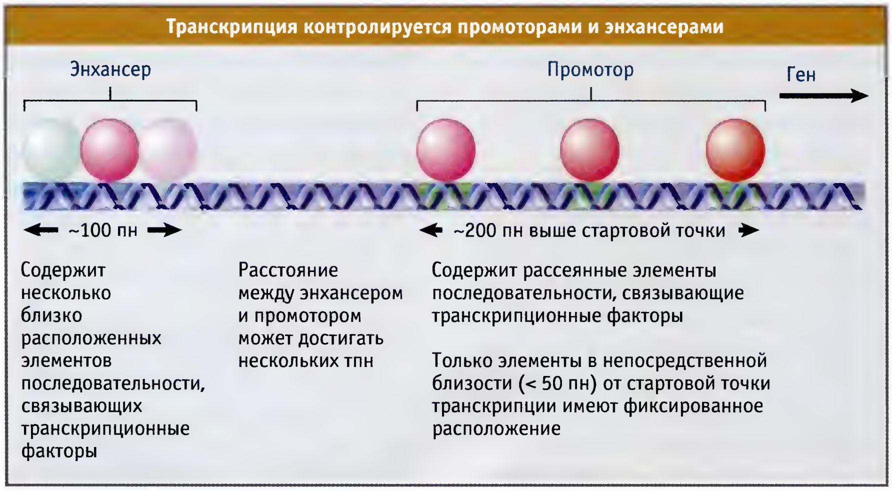


РИС. 24.1 Типичный ген, транскрибируемый РНК-полимеразой II, имеет промотор, расположенный выше сайта, с которого начинается транскрипция. Промотор несет несколько коротких (<10 пн) элементов, связывающих транскрипционные факторы; эти элементы расположены на участке длиной более 200 пн. Энхансер содержит набор более компактно расположенных элементов, которые тоже связывают транскрипционные факторы. Он может находиться в нескольких тысячах пар нуклеотидов от точки начала транскрипции. (ДНК может быть скручена или как-то иначе организована таким образом, что факторы транскрипции на промоторе и на энхансере взаимодействуют между собой, образуя крупный белковый комплекс)

экспрессируемые промоторы (соответствующие гены иногда называют «генами домашнего хозяйства») имеют вышележащие элементы в последовательности, распознаваемые часто встречающимися активаторами. Никакое сочетание элементов/факторов не является единственно возможным для данного промотора, указывая, что инициация РНК-полимеразой II может быть организована множеством различных способов. Промоторы, работающие только в определенное время или в определенном месте, несут элементы последовательности, требующие участия активаторов, доступных только в данное время и в данном месте.

Элементы последовательности промотора определяют функционально: они должны находиться в непосредственной близости от стартовой точки и должны обеспечивать инициацию транскрипции. **Энхансер** — еще один тип последовательности, участвующей в инициации. Его определяют как последовательность, стимулирующую инициацию, но расположенную на значительном удалении от стартовой точки. Энхансерные элементы часто представляют собой мишени для тканеспецифичной или временно́й регуляции. На рис. 24.1 показаны основные свойства промоторов и энхансеров.

Компоненты энхансера напоминают структуру промотора, поскольку тоже состоят из ряда модульных элементов, однако, как правило, более компактны. Энхансеры не обязательно должны быть расположены вблизи стартовой точки.

Белки, связывающиеся с элементами энхансеров, взаимодействуют с белками, связывающимися с элементами промотора. Различия между энхансерами и промоторами являются скорее функциональными, нежели фундаментальными. Такая точка зрения подтверждается фактом, что некоторые типы отдельных элементов встречаются как в промоторах, так и в энхансерах.

Транскрипция у эукариотических организмов имеет чаще всего положительную регуляцию: транскрипционный фактор для активации промотора или ряда промоторов, содержащих общую целевую последовательность, предоставляется под тканеспецифичным контролем. Регуляция транскрипции с помощью специфичной репрессии целевого промотора встречается реже.

Транскрипционная единица у эукариот обычно содержит один отдельный ген, а терминация происходит за пределами кодирующего региона. Терминация транскрипции в данном случае не имеет того регуляторного значения, которое она имеет у прокариот. РНК-полимеразы I и III терминируют на определенных последовательностях путем изученных реакций, однако способ терминации РНК-полимеразы II пока неясен. Создание 3′-конца зрелой мРНК, однако, не является следствием терминации, а является результатом реакции расщепления первичного транскрипта (см. гл. 26 «Сплайсинг и процессинг РНК»).

24.2 Эукариотическая ДНК-полимераза состоит из многих субъединиц

Ключевые понятия:

- РНК-полимераза I синтезирует рРНК в ядрышке.
- РНК-полимераза II синтезирует мРНК в нуклеоплазме.
- РНК-полимераза III синтезирует малые РНК в нуклеоплазме.
- Все эукариотические РНК-полимеразы имеют ~12 субъединиц и представляют собой комплексы массой >500 кДа.
- Некоторые субъединицы являются общими для всех трех РНК-полимераз.
- Самая крупная субъединица РНК-полимеразы II несет CTD (C-концевой домен), состоящий из множественных повторов гептамера.

Три эукариотические РНК-полимеразы имеют различное расположение в ядре в соответствии с генами, которые они транскрибируют.

Наиболее заметна активность фермента РНК-полимеразы I, которая располагается в ядрышке и обусловливает транскрипцию генов, кодирующих рРНК. Она отвечает за синтез большей части клеточных РНК (по массе).

РНК-полимераза II располагается в нуклеоплазме (содержимом ядра за исключением ядрышка). Она представляет собой бо́льшую часть оставшейся транскрипционной активности и отвечает за синтез гетерогенной ядерной РНК (гяРНК, hnРНК), предшественника мРНК.

РНК-полимераза III имеет сравнительно невысокую ферментативную активность. Этот нуклеоплазматический фермент синтезирует тРНК и другие малые РНК.

Все эукариотические РНК-полимеразы являются крупными белками, действующими в виде агрегатов весом более 500 кДа. Обычно они содержат порядка 12 субъединиц. Очищенный фермент способен выполнять транскрипцию РНК на матрице, однако не способен избирательно инициировать транскрипцию на промоторах. Общий состав эукариотической РНК-полимеразы II (на примере фермента из *Saccaromyces cerevisiae*) показан на рис. 24.2. Две самых крупных субъединицы гомологичны β- и β′-субъединицам бактериальной РНК-полимеразы. Три других субъединицы общие для всех эукариотических РНК-полимераз, таким образом, они являются компонентами также и РНК-полимераз I и III.

Самая крупная субъединица РНК-полимеразы II имеет **C-концевой домен** (carboxy-terminal domain, CTD), состоящий из множества повторов консенсусной последовательности из семи аминокислотных остатков. Эта последовательность уникальна для РНК-полимеразы II. У дрожжей содержится около 26 таких повторов, а у млекопитающих — около 50. Число повторов весьма важно, поскольку делеции с удалением более половины повторов летальны (у дрожжей). CTD может быть фосфорилирован по сериновым и треониновым остаткам; это используется в реакциях инициации транскрипции (см. разд. 24.11 «Инициация сопровождается освобождением промотора»).

РНК-полимеразы митохондрий и хлоропластов меньше по размеру и, скорее, напоминают бактериальные РНК-полимеразы, нежели ядерные ферменты.

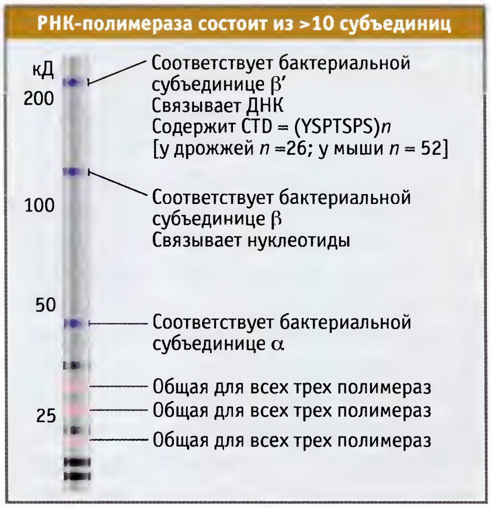


РИС. 24.2 Некоторые субъединицы являются общими для всех классов эукариотических РНК-полимераз, а некоторые гомологичны бактериальным РНК-полимеразам

Поскольку геномы органелл меньше, резидентная полимераза транскрибирует относительно немного генов, и контроль траснкрипции, по-видимому, здесь значительно проще (если вообще существует). Таким образом, ферменты органелл в чем-то аналогичны фаговым ферментам, у которых отсутствует способность реагировать на более сложное окружение.

Различие между эукариотическими РНК-полимеразами вытекает из их ответа на бициклический октапептид α-аманитин. Практически во всех эукариотических клетках активность РНК-полимеразы II быстро ингибируется низкими концентрациями α-аманитина. Однако РНК-полимераза I на него не реагирует. Ответ РНК-полимеразы III на α-аманитин менее консервативен; в животных клетках активность фермента ингибируется при высокой концентрации этого пептида, а в клетках дрожжей и насекомых не ингибируется.

24.3 Промоторные элементы определяют с помощью мутагенеза и футпринтинга

Ключевые понятия:

- Промоторы выявляют по их способности вызывать транскрипцию присоединенной к ним последовательности в соответствующей тест-системе *in vitro* и *in vivo*.

Первым этапом характеризации промотора является определение общей длины ДНК, содержащей все необходимые элементы. Чтобы это сделать, нужно иметь тест-систему, где промотор отвечает за продукцию белка, который легко проанализировать. Исторически использовали несколько типов таких систем:

- В *системе с использованием ооцитов Xenopus laevis* матричную ДНК инъецируют в ядро ооцита. РНК-транскрипт можно выделить и проанализировать. Основной недостаток этой системы в том, что она ограничена условиями, которые имеются в ооцитах. Система позволяет охарактеризовать последовательности ДНК, но не факторы, которые ее связывают.
- *Системы трансфекции* позволяют ввести экзогенную ДНК в культуру клеток и экспрессировать ее. Эти системы являются системами *in vivo* в том смысле, что транскрипция выполняется тем же аппаратом, который отвечает за экспрессию собственного генома клетки. Однако в эксперименте матрица может содержать такие гены, которые в норме обычно не транскрибируются в хозяйской клетке. Использование этих систем можно расширить, выбирая различные клетки-хозяева.
- *Трансгенные системы* представляют собой добавление гена в зародышевую линию клеток животного. За экспрессией **трансгена** можно наблюдать в какой-либо (или во всех сразу) тканях животного. Трансгенные и трансфекционные системы обладают общим недостатком: дополнительный ген часто присутствует во множестве копий и интегрирован в различных положениях в геноме. Различия между экспрессией гена *in vitro* и его экспрессией в качестве трансгена могут дать важную информацию о роли геномного контекста для данного гена.
- Система *in vitro* использует классический подход: очищение всех компонентов и манипулирование условиями до тех пор, пока не будет достигнута правильная инициация. «Правильная» инициация определяется как продукция РНК, которая начинается с соответствующего 5′-концу мРНК сайта (или предшественников рРНК или тРНК). В итоге это позволяет охарактеризовать конкретные элементы последовательности в промоторе и транскрипционные факторы, их связывающие.

При анализе промотора важно, чтобы изменялась *только* последовательность промотора. На РИС. 24.3 показано, что выше промотора всегда помещают единую протяженную последовательность, чтобы иметь гарантию, что он всегда работает в одном и том же контексте. В системах *in vitro* терминация не происходит правильным образом, поэтому матрицу обрезают на некотором расстоянии от промотора (обычно примерно на 500 пн ниже). Это дает гарантию, что все полимеразы «соскакивают» в одной и той же точке, давая легко распознаваемый транскрипт.

Можно начать с определенного фрагмента ДНК, который способен инициировать транскрипцию в одной из этих систем. Границы последовательности, представляющей собой промотор, можно определить путем уменьшения длины этого фрагмента с любой стороны до тех пор, пока в некоторой точке она не перестанет быть активной, как показано на РИС. 24.4. Начало можно определить, постепенно удаляя материал

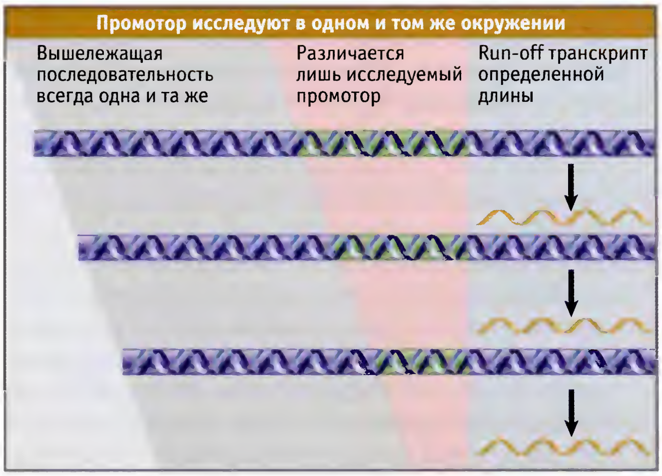


РИС. 24.3 Промотор исследуют путем модификации последовательности, заключенной между константной последовательностью выше и константным участком для транскрипции ниже

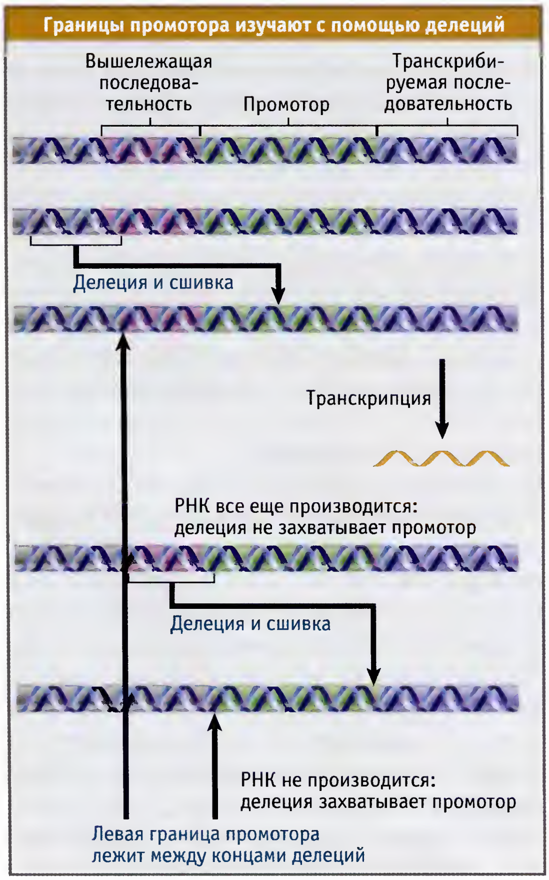


РИС. 24.4 Границы промотора можно определить путем делеций, которые постепенно удаляют материал с одной стороны. Если предыдущая делеция не предотвращает синтез РНК, а в результате следующей транскрипция прекращается, граница промотора лежит между этими делециями

ниже промотора до тех пор, пока он не потеряет свою функцию. Чтобы определить начало промотора, необходимо воссоединить укороченную промоторную последовательность с транскрибируемой последовательностью (поскольку в противном случае не будет продукта, который можно анализировать).

Когда границы промотора определены, значение каждого основания в нем может быть изучено путем внесения точечных мутаций или иной реорганизации последовательности. Как и в случае с бактериальной РНК-полимеразой, эти мутации можно назвать *повышающими* и *понижающими*. Некоторые из реорганизаций влияют только на *скорость* инициации; другие влияют на *сайт* инициации, что видно по изменению стартовой точки. Чтобы быть уверенными, что мы имеем дело со сравнимыми продуктами, в каждом случае необходимо характеризовать 5′-конец РНК.

Можно применить несколько критериев для идентификации компонентов последовательности промотора (или любого другого участка ДНК):

- Введение мутаций и тестирование функции *in vitro* и *in vivo*. (Сейчас существует множество технологий внесения точечных мутаций в определенные пары оснований и в принципе каждую позицию промотора можно мутировать и протестировать такую мутантную последовательность *in vitro* и *in vivo*.)
- Футпринтинг белков, действующих путем связывания с сайтом. Должна существовать корреляция между способностью мутаций предотвращать функцию промотора и предотвращать связывание фактора.
- Если некий сайт, распознаваемый определенным фактором, присутствует в нескольких промоторах, можно вывести консенсусную последовательность, с которой будет связываться этот фактор. И новый промотор должен отвечать на фактор, если в него внедрить соответствующую копию консенсусного элемента.

24.4 Промотор РНК-полимеразы I состоит из двух частей

Ключевые понятия:

- Промотор РНК-полимеразы I состоит из центральной (core) области и контрольного элемента, расположенного выше (UPE, upstream control element).
- Фактор UBF1 накручивает ДНК вокруг белковой структуры, чтобы сблизить центральную область промотора и UPE.
- SL1 включает в себя фактор TBP, который используется при инициации всеми тремя РНК-полимеразами.
- РНК-полимераза связывается с комплексом UBF1-SL1 в центральной области промотора.

РНК-полимераза I осуществляет транскрипцию с единственного типа промоторов (присущего генам рибосомальной РНК). Транскрипт гена РНК включает в себя последовательности и большой, и малой рРНК, кото-

рые впоследствии высвобождаются путем расщепления и процессинга. Гены рРНК существуют во множестве копий, которые чередуются с нетранскрибируемыми **спейсерами** и организованы в кластеры, как обсуждалось в разд. 6.8 «Гены рРНК формируют тандемные повторы». Структурная организация промотора и события, происходящие при инициации, показаны на рис. 24.5.

Промотор состоит из двух отдельных районов. **Ядро промотора** окружает стартовую точку, находясь от −45 до +20, и его достаточно для инициации транскрипции. Обычно это G-C-богатый участок (что нетипично для промоторов), за исключением единственного консервативного элемента — короткой A-T-богатой последовательности вокруг стартовой точки, называемой Inr. Эффективность ядра промотора, однако, существенно повышается вышележащим промоторным элементом (UPE, upstream promotor element). UPE — это еще одна G-C-богатая последовательность, похожая на ядро промотора, расположенная на участке от −180 до −107. Такой тип организации обычен для промоторов pol I у многих видов, хотя конкретная последовательность может значительно различаться.

РНК-полимеразе I для работы необходимы два дополнительных фактора. Фактор, связывающий ядро промотора, состоит из четырех белков (его называют SL1, TIF-BP или Rib1 у различных видов.) Один из этих белков, TATA-связывающий белок (TBP, TATA-binding protein), также необходим для инициации РНК-полимеразами II и III (см. разд. 24.8 «TBP — универсальный фактор»). TBP не связывается непосредственно с G-C-богатым участком ДНК, и его посадка на ДНК осуществляется с помощью других компонентов фактора, связывающегося с ядром промотора. Вполне вероятно, что TBP взаимодействует с РНК-полимеразой (по-видимому, с общей субъединицей или с каким-то белком, который консервативен у полимераз). Кор-связывающий фактор дает РНК-полимеразе I возможность инициировать на промоторе с низкой, но постоянной частотой.

Кор-связывающий фактор в основном отвечает за то, чтобы РНК-полимераза располагалась на стартовой точке правильным образом. Вскоре мы увидим, что примерно такую же функцию для РНК-полимераз II и III выполняет фактор, состоящий из TBP в комплексе с другими белками. Таким образом, общей чертой при инициации всеми тремя полимеразами является зависимость от «позиционирующих» факторов, состоящих из TBP и ассоциированных с ним белков, специфичных к каждому определенному типу промоторов.

Для инициации с высокой частотой необходимо участие дополнительного фактора UBF. Это отдельный полипептид, связывающий с G-C-богатым элементом на UPE. Одним из свидетельств взаимодействия UBF с кор-связывающим фактором является важное значение расстояния между UPE и ядром промотора. Это расстояние можно изменить, включив целое число витков ДНК. UBF связывается с малой бороздкой ДНК и закручивает

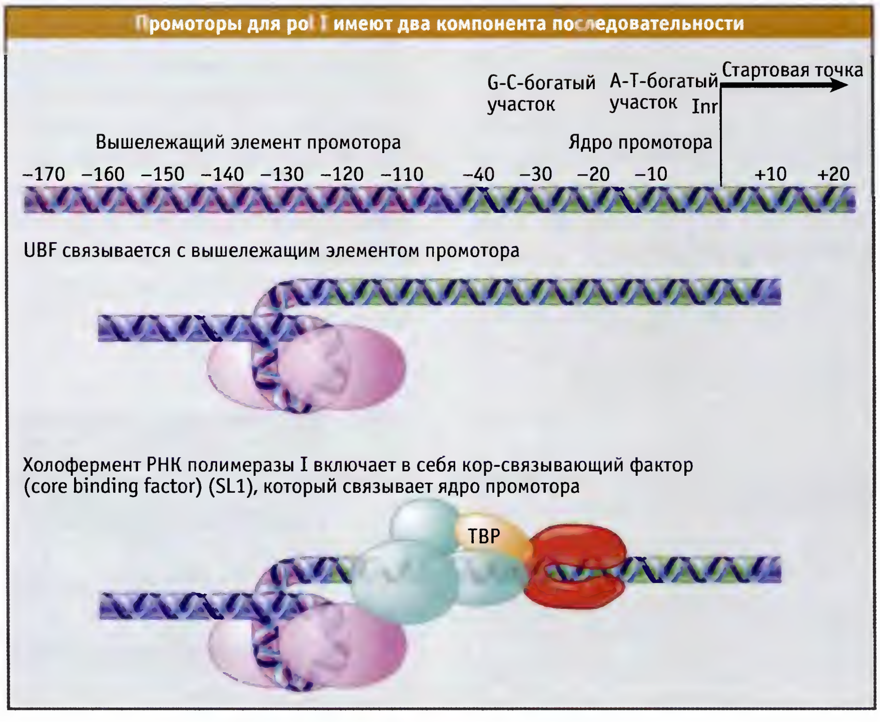


Рис. 24.5 Траснкрипционная единица для РНК-полимеразы I включает в себя ядро промотора, удаленное от вышележащего промоторного элемента примерно на 70 пн. Связывание UBF с UPE увеличивает способность кор-связывающего фактора соединяться с ядром промотора. Кор-связывающий фактор (SL1) позиционирует РНК-полимеразу I на стартовой точке

ДНК в петлю почти на 360°, что приближает ядро и UPE друг к другу.

На рис. 24.5 показана инициация транскрипции как серия последовательных взаимодействий. РНК-полимераза I, однако, существует как холофермент, который содержит большинство или даже все факторы, необходимые для инициации, и который, по-видимому, рекрутируется непосредственно на промотор.

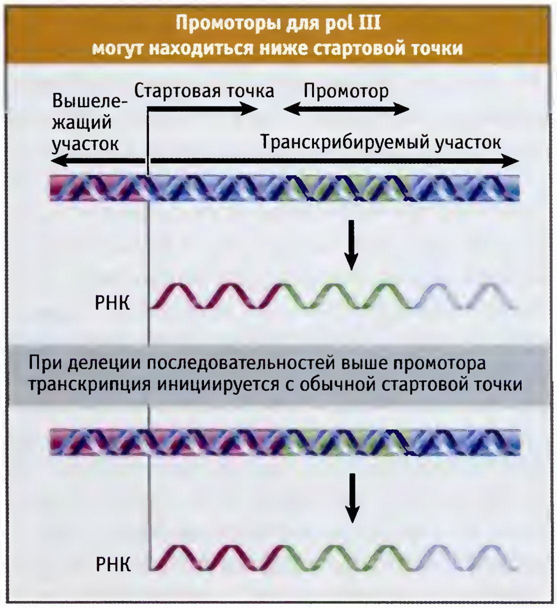


РИС. 24.6 **Делеционный анализ показывает, что промотор для 5S РНК – внутренний; инициация происходит на фиксированном расстоянии (около 55 пн) выше промотора**

24.5 РНК-полимераза III использует и вышележащие, и нижележащие промоторы

Ключевые понятия:

- РНК-полимераза III узнает два типа промоторов.
- Внутренние промоторы имеют короткие консенсусные последовательности, расположенные в транскрипционной единице, и вызывают инициацию на определенном расстоянии выше промотора.
- Вышележащие промоторы содержат три коротких консенсусных последовательности, связывающие транскрипционные факторы.

Распознавание промоторов РНК-полимеразой III иллюстрирует относительность ролей факторов транскрипции и собственно фермента. Промоторы можно разделить на два основных класса, распознаваемых по-разному и различными группами факторов. Промоторы для генов 5S и тРНК — *внутренние*; они находятся ниже стартовой точки транскрипции. Промоторы для генов мяРНК (малых ядерных РНК; от *англ.* small nuclear, snРНК — РНК) находятся выше точки начала транскрипции (как и другие обычные промоторы). В обоих случаях конкретные элементы, необходимые для работы промоторов, состоят исключительно из последовательностей, распознаваемых факторами транскрипции, которые, по сути, управляют связыванием РНК-полимеразы.

До момента определения промотора генов 5S у *Xenopus laevis* все попытки установить промоторную последовательность предполагали, что она должна находиться выше стартовой точки. Однако делеционный анализ показал, что продукт 5S РНК продолжает синтезироваться, даже когда удалена вся последовательность выше гена!

Когда делеции продолжаются в гене, продукт близкого к обычной 5S РНК размера продолжает синтезироваться до тех пор, пока делеция не затронет основание +55. На РИС. 24.6 показано, что первая часть РНК-продукта при этом соответствует плазмидной ДНК, а вторая часть представляет собой сегмент, оставшийся от обычной последовательности 5S РНК. При этом когда делеция продолжается за +55, транскрипция прекращается. Следовательно, промотор находится *ниже позиции* +55, но заставляет РНК-полимеразу III инициировать транскрипцию на более или менее фиксированном расстоянии выше него.

Когда делеции производятся в гене с его дистального конца, транскрипция не изменяется, пока первые 80 пн остаются интактными. Как только делеции затрагивают этот район, транскрипция прекращается. Это указывает на нижнюю границу промотора в положении примерно +80.

Таким образом, промотор для транскрипции 5S РНК находится между позициями +55 и +80 внутри гена. Фрагмент последовательности, содержащий этот район, может вызывать инициацию транскрипции на любой ДНК, в которую он помещен, со стартовой точки примерно на 55 пн выше. (Стартовая точка дикого типа уникальна; в тех же делециях, где она отсутствует, транскрипция начинается с пуринового основания, ближайшего к точке в 55 пн выше промотора.)

Структура трех типов промоторов для РНК-полимеразы III схематично представлена на РИС. 24.7. Существует два типа внутренних промоторов. Каждый имеет

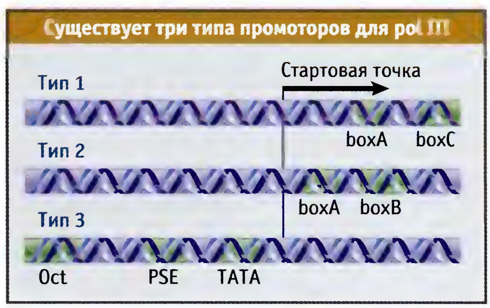


РИС. 24.7 **Промоторы для РНК-полимеразы III могут состоять из двух частей последовательности ниже стартовой точки, где boxA отделен от boxC или boxB, или могут состоять из отдельных последовательностей выше стартовой точки (Oct, PSE, TATA)**

структуру из двух частей, в которой короткие элементы последовательности разделены вариабельной последовательностью. Тип 1 состоит из последовательности boxA, отделенной от последовательности boxC, а тип 2 состоит из последовательности boxA, отделенной от boxB. Расстояние между boxA и boxB в промоторе типа 2 может сильно варьировать, однако эти участки не могут располагаться слишком близко друг к другу без потери функции промотора. Широко распространенная группа промоторов типа 3 имеет три элемента последовательности, расположенных выше стартовой точки.

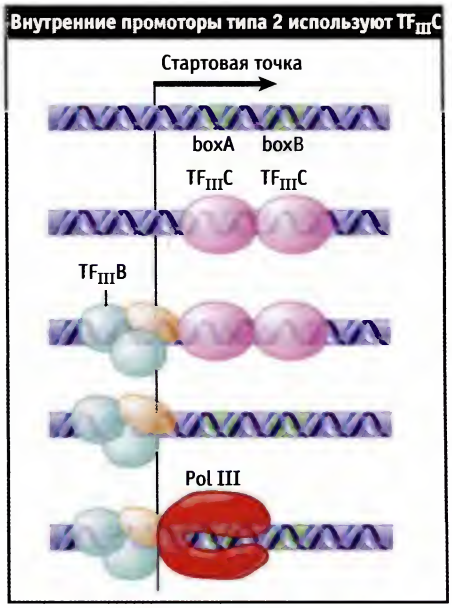


РИС. 24.8 На внутренних промоторах типа 2 для pol III используется связывание $TF_{III}C$ с последовательностями boxA и boxB для рекрутирования позиционного фактора $TF_{III}B$, который, в свою очередь, рекрутирует РНК-полимеразу III

24.6 $TF_{III}B$ — обязательный фактор для промоторов РНК-полимеразы III

Ключевые понятия:

- **$TF_{III}A$ и $TF_{III}C$ связываются с консенсусными последовательностями и способствуют связыванию $TF_{III}B$ со стартовой точкой.**
- **$TF_{III}B$ несет TBP в качестве одной из субъединиц и делает возможным связывание РНК-полимеразы.**

Если рассматривать процесс инициации в деталях, то взаимодействие факторов на двух типах внутренних промоторов отличается, однако в целом оно очень схоже. $TF_{III}C$ связывается ниже стартовой точки, либо независимо (промоторы тип 2), либо совместно с $TF_{III}A$ (промоторы тип 1). Присутствие $TF_{III}C$ позволяет позиционирующему фактору $TF_{III}B$ связаться со стартовой точкой. Затем рекрутируется РНК-полимераза.

На РИС. 24.8 представлены стадии инициации на внутренних промоторах типа 2. $TF_{III}C$ связывается с boxA и boxB. Это дает возможность $TF_{III}B$ связаться со стартовой точкой. На этой стадии уже может связываться РНК-полимераза III.

Отличие внутреннего промотора типа 1 в том, что $TF_{III}A$ должен связаться с boxA, чтобы дать возможность $TF_{III}C$ связаться с boxC. На РИС. 24.9 показано, что, как только $TF_{III}C$ связывается, дальнейшие события разворачиваются так же, как и на промоторах типа 2, со связыванием со стартовой точкой $TF_{III}B$ и вовлечением в комплекс РНК-полимеразы III. Промоторы типа 1 найдены только в генах для 5S РНК.

$TF_{III}A$ и $TF_{III}C$ — **факторы агрегации**, чья единственная роль состоит в том, чтобы помочь $TF_{III}B$ связаться с нужным участком. Как только $TF_{III}B$ свяжется с ДНК, $TF_{III}A$ и $TF_{III}C$ можно удалить с промотора (высокой концентрацией соли *in vitro*) без какого-либо влияния на инициацию транскрипции. *$TF_{III}B$ остается связанным в районе стартовой точки, и его присутствия достаточно, чтобы РНК-полимераза III связалась с промотором.* Следовательно, $TF_{III}B$ — единственный истинный фактор инициации, необходимый РНК-полимеразе III. Такая последовательность событий объясняет, каким

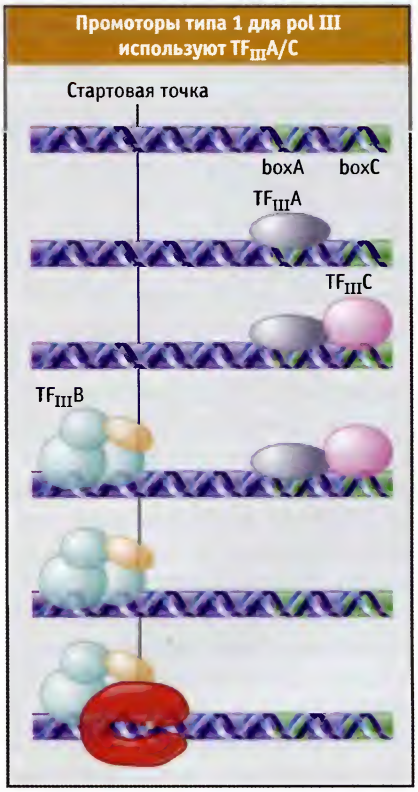


РИС. 24.9 На внутренних промоторах типа 1 для pol III используются факторы агрегации $TF_{III}A$ и $TF_{III}C$, которые связываются с последовательностями boxA и boxC для рекрутирования позиционного фактора $TF_{III}B$, который рекрутирует РНК-полимеразу III

образом промоторные участки (boxes) ниже стартовой точки заставляют РНК-полимеразу связаться со стартовой точкой значительно выше. Несмотря на то что способность транскрибировать эти гены реализуется с помощью внутренних промоторов, изменения участка сразу выше стартовой точки могут влиять на эффективность транскрипции.

$TF_{III}C$ — это крупный белковый комплекс (более 500 кДа), сравнимый по размеру с самой РНК-полимеразой, содержащий шесть субъединиц. $TF_{III}A$ принадлежит к белкам, имеющим ДНК-связывающий мотив, называемый «цинковым пальцем» (см. разд. 25.9 «Цинковые пальцы» связывают ДНК»). Позиционный фактор $TF_{III}B$ состоит из трех субъединиц. Он включает в себя тот же белок ТВР, который присутствует в кор-связывающем факторе для промоторов pol I, а также в соответствующем транскрипционном факторе ($TF_{II}D$) для РНК-полимеразы II. Также он содержит Brf, похожий на фактор РНК-полимеразы II $TF_{II}B$. Третью субъединицу называют B″; если ДНК-дуплекс частично расплавлен, ее участие необязательно, что указывает на ее функцию — образование «глазка» в процессе инициации транскрипции. Роль B″ можно сравнить с ролью сигма-фактора бактериальной РНК-полимеразы (см. разд. 11.16 «Инициация на данном промоторе зависит от типа сигма-фактора»).

Вышележащий промотор играет обычную роль для РНК-полимеразы III. На примере, показанном на рис. 24.8, можно видеть три вышележащих элемента. Эти элементы также имеются в промоторах генов мяРНК, транскрибируемых РНК-полимеразой II. (Гены некоторых мяРНК транскрибируются РНК-полимеразой II, а некоторых — РНК-полимеразой III.) Вышележащие элементы в составе промоторов действуют сходным образом для полимераз II и III.

Инициация на вышерасположенном промоторе РНК-полимеразы III может происходить на коротком участке, непосредственно предшествующем стартовой точке и несущем лишь TATA-элемент. Однако эффективность транскрипции значительно возрастает в присутствии элементов PSE и Oct. Факторы, связывающиеся с этими элементами, действуют совместно. (Элемент PSE может быть незаменимым для промоторов РНК-полимеразы II, но является дополнительным для промоторов РНК-полимеразы III; его название происходит от сокращения словосочетания proximal sequence element).

TATA-элемент обусловливает тип полимеразы (II или III), которая распознает промотор мяРНК. Этот элемент связывается фактором, включающим в себя ТВР, который и распознает последовательность ДНК. ТВР ассоциирован с другими белками, специфичными к типу промотора. Функция ТВР и связанных с ним белков состоит в правильном позиционировании РНК-полимеразы на стартовой точке. Мы обсудим это более детально для РНК-полимеразы II (см. разд. 24.8 «ТВР — универсальный фактор»).

Для обоих типов промоторов РНК-полимеразы III факторы действуют одинаково. *Факторы связываются с промотором до того, как РНК-полимераза узнает промотор.* Они образуют **преинициаторный комплекс**, управляющий связыванием РНК-полимеразы. РНК-полимераза III сама по себе не распознает промоторную последовательность, однако садится рядом с факторами, которые связываются чуть выше стартовой точки. Для внутренних промоторов типа 1 и типа 2 факторы агрегации обеспечивают связывание $TF_{III}B$ (который включает и ТВР) с последовательностью чуть выше стартовой точки, таким образом давая информацию о ее местонахождении. Для вышележащих промоторов $TF_{III}B$ связывается напрямую с участком, включающим TATA-бокс. Вышесказанное означает, что безотносительно расположению промоторных последовательностей, фактор(ы) связываются вблизи стартовой точки в таком порядке, чтобы управлять посадкой РНК-полимеразы III.

24.7 Стартовая точка для РНК-полимеразы II

Ключевые понятия:

- Для инициации транскрипции РНК-полимеразе II необходимо присутствие главных транскрипционных факторов (называемых $TF_{II}X$).
- Промоторы РНК-полимеразы II имеют короткую консервативную последовательность Py_2CAPy_5 (инициатор InR) в стартовой точке.
- TATA-бокс — общий компонент промоторов для РНК-полимеразы II, состоящий из A-T-богатого октамера, расположенного на 25 пн выше стартовой точки.
- DPE — общий компонент для промоторов РНК-полимеразы II, не несущих TATA-бокса.
- Главный элемент промотора для РНК-полимеразы II включает в себя InR и либо TATA-бокс, либо DPE.

Основные принципы организации аппарата транскрипции генов, кодирующих белки, были установлены при обнаружении того факта, что очищенная РНК-полимераза II может катализировать синтез мРНК, но не может инициировать транскрипцию без добавления еще одного экстракта. Фракционирование этого экстракта привело к выделению основных факторов транскрипции — группы белков, необходимых для инициации транскрипции РНК-полимеразой II на всех типах промоторов. РНК-полимераза II в комплексе с этими факторами представляет собой базовый транскрипционный аппарат, необходимый для траскрипции всех промоторов. Эти главные факторы называют $\mathbf{TF_{II}X}$, где «X» — буква-идентификатор собственно фактора. Субъединицы РНК-полимеразы II и главные факторы транскрипции у эукариот консервативны.

Точкой отсчета для изучения организации промотора для нас является определение ядра промотора как самой короткой последовательности, на которой РНК-полимераза II может инициировать транскрипцию. В принципе ядро промотора может работать в любой клетке. Оно заключает в себе минимальную последовательность, позволяющую факторам транскрипции собраться

на стартовой точке. Ядра промоторов участвуют в механизме связывания с ДНК и позволяют РНК-полимеразе II инициировать транскрипцию. Однако ядро промотора выполняет свою функцию с низкой эффективностью. Для адекватного уровня эффективности необходимы белки, называемые активаторами (см. разд. 24.13 «Активаторы связываются с короткими элементами последовательности»). Активаторы имеют случайные названия, отражающие историю их открытия.

Можно ожидать, что компоненты последовательности, участвующие в связывании РНК-полимеразы и главных факторов транскрипции, будут консервативны у большинства или даже у всех промоторов. Что касается промоторов бактерий, то при сравнении промоторов для РНК-полимеразы II сходство участка возле стартовой точки ограничена довольно короткими последовательностями. Эти элементы соответствуют последовательностям, заключающим в себе собственно промоторную функцию. На РИС. 24.10 схематично представлена конструкция типичного ядра промотора для pol II.

В точке начала транскрипции нет протяженного сходства последовательностей, однако наиболее часто в качестве первого основания мРНК встречается А, с обеих сторон фланкированный пиримидинами. (Это описание также применимо к стартовой последовательности САТ бактериальных промоторов.) Этот участок называют *инициатором* (**Inr**), и его можно в общем описать формулой Py_2CAPy_5. Inr находится между позициями −3 и +5.

Многие промоторы имеют последовательность, называемую **TATA-бокс**, у высших эукариот расположенную обычно примерно в 25 пн выше стартовой точки, и являются единственным вышележащим промоторным элементом, имещим достаточно фиксированное расположение относительно стартовой точки. Последовательность ядра TATA-бокса TATAA далее сопровождается обычно еще тремя A-T парами. TATA-бокс часто находится в окружении G-C-богатых последовательностей, которые могут быть одним из факторов его функционирования. Он почти идентичен с последовательностью −10 бактериальных промоторов; фактически, они очень похожи, за исключением расположения в районе −25 вместо −10.

Замены отдельных оснований в TATA-боксе действуют как мутации, очень сильно ослабляющие функцию. Некоторые такие мутации меняют лишь ориентацию A-T пар; таким образом, состав последовательности сам по себе еще не определяет ее функцию. Следовательно, TATA-бокс представляет собой элемент, поведение которого аналогично нашей концепции бактериального промотора: это определенная короткая последовательность непосредственно выше стартовой точки, необходимая для транскрипции.

Существуют промоторы, **не имеющие TATA-бокса**. Исследования промоторных последовательностей показывают, что 50% или даже более промоторов могут не иметь TATA-бокса. Когда промотор не содержит TATA-бокса, он обычно имеет другой элемент, DPE (downstream promotor element), расположенный между +28 и +32.

Большинство ядер промоторов состоят либо из TATA-бокса плюс Inr, либо Inr плюс DPE.



РИС. 24.10 Минимальный промотор для pol II несет TATA-бокс в точке примерно 25 пн выше Inr. TATA-бокс имеет консенсусную последовательность TATAA. Inr содержит пиримидиновые основания (Y), окружающие CA в стартовой точке. DPE находится ниже стартовой точки. Последовательность показана на кодирующей цепи

24.8 TBP — универсальный фактор

Ключевые понятия:

- TBP — компонент направляющего фактора, необходимого для связывания каждого типа РНК-полимеразы с ее промотором.
- Этим фактором для РНК-полимеразы II является $TF_{II}D$, состоящий из TBP и 11 TAF, общей массой ~800 кДа.

Первым шагом формирования комплекса на промоторе, содержащем TATA-бокс, является связывание фактора **$TF_{II}D$** с участком, продолжающимся выше TATA-последовательности. $TF_{II}D$ содержит два типа компонентов. Распознавание TATA-бокса обеспечивается **TATA-связывающим белком** (**TBP**, TATA-binding protein), небольшим белком с массой около 30 кДа. Другие субъединицы называют **TAF** (от *англ.* TBP-associated factors — TBP-ассоциированные факторы). Некоторые TAF находятся в стехиометрическом отношении к TBP; другие встречаются в меньших количествах. $TF_{II}D$, содержащие различные TAF, могут распознавать различные промоторы. Некоторые (субстехиометрические) TAF являются тканеспецифичными. Общая масса $TF_{II}D$ составляет обычно около 800 кДа. В таком $TF_{II}D$ содержатся TBP и 11 TAF, чья масса варьирует от 30 до 250 кДа. Названия TAF в составе $TF_{II}D$ образованы в форме $TAF_{II}00$, где «00» представляет собой молекулярную массу субъединицы.

Позиционирующие факторы, состоящие из TBP, ассоциированного с определенным набором TAF, отве-

чают за узнавание всех классов промоторов. $TF_{III}B$ (для промоторов pol III) и SL1 (для промоторов pol I) могут рассматриваться как состоящие из TBP, ассоциированного с особой группой белков — заместителей TAF, присутствующих в $TF_{II}D$. TBP — это основной компонент, встроенный в каждый тип промотора с помощью различных механизмов. В случае промоторов для РНК-полимеразы II основная особенность позиционирования заключается в фиксированном расстоянии от стартовой точки до TATA-бокса.

На РИС. 24.11 показано, что в каждом случае позиционирующие факторы распознают промотор по-разному. На промоторах для РНК-полимеразы III $TF_{III}B$ связывается рядом с $TF_{III}C$. На промоторах для РНК-полимеразы I SL1 связывается в комплексе с UBF. И лишь $TF_{II}D$ в одиночку отвечает за распознавание промоторов для РНК-полимеразы II. На промоторе, имеющем TATA-элемент, TBP самостоятельно связывается с ДНК, однако на других промоторах он может быть вовлечен в реакцию в ассоциации с другими белками, узнающими определенные последовательности. Какими бы способами он ни вовлекался в инициаторный комплекс, главная цель процесса состоит во взаимодействии с РНК-полимеразой.

$TF_{II}D$ широко распространен, однако не уникален. Все многоклеточные эукариоты экспрессируют также альтернативный комплекс, в котором есть TLF (TBP-like factor, TBP-подобный фактор) вместо TBP. TLF обычно примерно на 60% схож с TBP. По-видимому, он инициирует образование комплекса с помощью обычного набора TF_{II}-факторов. Однако TLF не связывается с TATA-боксом, и мы пока не знаем, как он работает. У *Drosophila* есть также и третий фактор, TRF1, действующий подобно TBP и связывающий свой набор TAF для образования комплекса, который функционирует как альтернатива $TF_{II}D$ на определенных промоторах.

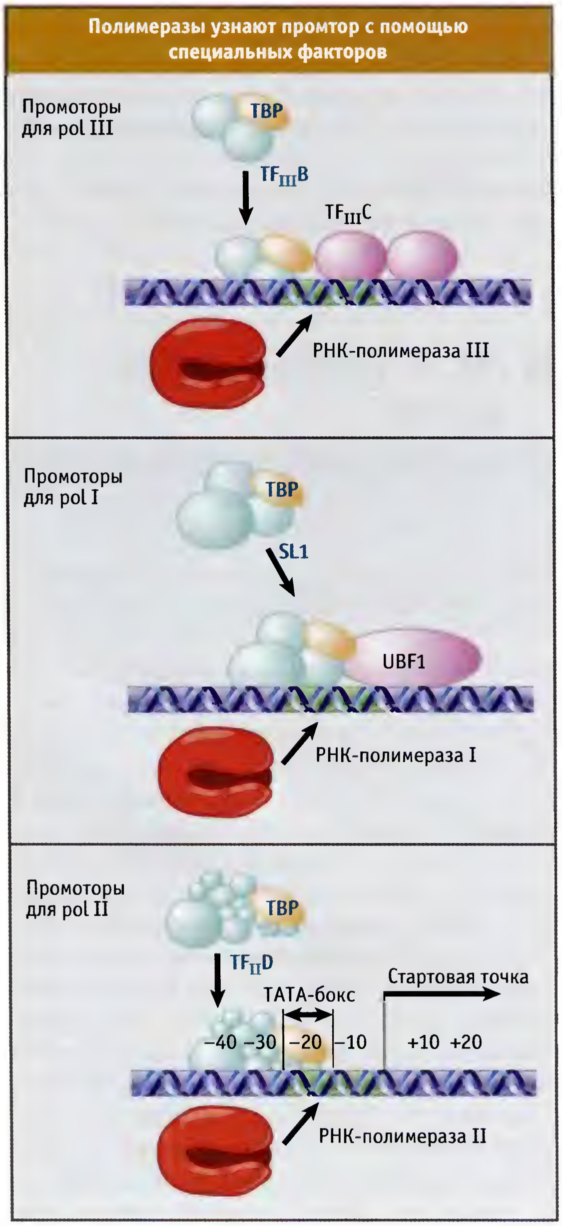


РИС. 24.11 РНК-полимеразы позиционируются на всех типах промоторов с помощью фактора, содержащего TBP

24.9 TBP связывает ДНК необычным способом

Ключевые понятия:

- TBP связывается с TATA-боксом в малой бороздке ДНК.
- Он полуокружает ДНК и изгибает ее примерно на 80°
- Некоторые из TAF напоминают гистоны и могут образовывать структуры, напоминающие гистоновый октамер (нуклеосому).

TBP имеет необычное свойство — он связывается с ДНК в малой бороздке (почти все известные ДНК-связывающие белки связываются с большой бороздкой). На кристаллической структуре TBP можно хорошо увидеть способ его связывания с ДНК. На РИС. 24.12 видно, что он окружает одну сторону ДНК, образуя полукруг вокруг двойной спирали. Внутренняя поверхность TBP связывается с ДНК, а его бо́льшая внешняя поверхность остается доступной для контактов с другими белками. ДНК-связывающий сайт состоит из C-концевого консервативного домена, а вариабельный N-конец доступен для взаимодействий с другими белками. Степень консервативности механизма инициации транскрипции можно оценить по такому факту: последовательность ДНК-связывающего TBP у дрожжей и человека имеет 80% совпадающих аминокислот.

В присутствии нуклеосом TBP связывается нестабильно. Нуклеосомы вызывают перемещение A-T-богатых последовательностей в малую бороздку; в результате они могут препятствовать связыванию TBP. Этот эффект может объяснить, почему присутствие нуклеосом предотвращает инициацию транскрипции.



РИС. 24.12 На поперечном разрезе видно, что TBP окружает ДНК со стороны малой бороздки. TBP состоит из двух родственных (на 40% идентичных) консервативных доменов, которые показаны светло- и темно-голубым цветом. Его N-конец имеет значительное разнообразие, показан зеленым цветом. Две цепи двойной спирали ДНК показаны светло- и темно-серым цветом. Фотография любезно предоставлена Stephen K. Burtley

TBP связывается с малой бороздкой и изгибает ДНК примерно на 80°, как показано на РИС. 24.13. TATA-бокс сгибается в направлении большой бороздки, расширяя малую бороздку. Деформация ДНК ограничена 8 пн TATA-бокса; с каждой стороны этой последовательности малая бороздка имеет свою обычную ширину около 5 Å, а в центре последовательности ширина бороздки более 9 Å. Это нарушает структуру ДНК, однако цепи еще не разделяются, поскольку спаривание оснований сохранено. Степень изгиба может варьировать в зависимости от конкретной последовательности TATA-бокса, и это коррелирует с эффективностью промотора.

Из такой структуры вытекают несколько функционально важных следствий. Изменение пространственной организации ДНК с обеих сторон TATA-бокса позволяет транскрипционным факторам и РНК-полимеразе плотнее соединиться друг с другом, чем это было бы возможно на линейной ДНК. Изгиб TATA-бокса соответствует раскручиванию ДНК примерно на одну треть витка, что компенсируется сверхспирализацией.

Присутствие TBP в малой бороздке, а также его связывание с другими белками в малой бороздке создают высокую плотность контактов белок-ДНК в этом районе. Связывание TBP с ДНК *in vitro* защищает примерно один виток двойной спирали на TATA-боксе, обычно на участке между –37 и –25. Связывание комплекса $TF_{II}D$ обычно защищает район уже от –45 до –10, и продолжается даже выше стартовой точки. При этом TBP — единственный транскрипционный фактор, связывающийся с ДНК специфично.

В свободном белковом комплексе $TF_{II}D$ фактор $TAF_{II}230$ связывается с TBP, занимая на нем вогнутую ДНК-связывающую поверхность. Фактически структура связывающего сайта, находящегося в N-концевом домене $TAF_{II}230$, имитирует поверхность малой бороздки ДНК. Эта молекулярная мимикрия позволяет $TAF_{II}230$ контролировать способность TBP связываться с ДНК; N-концевой домен $TAF_{II}230$ должен быть вытеснен с ДНК-связывающей поверхности TBP для того, чтобы $TF_{II}D$ связался с ДНК.

Некоторые TAF напоминают гистоны; в частности, $TAF_{II}42$ и $TAF_{II}62$ отдаленно схожи с гистонами H3 и H4, и образуют гетеродимер, используя такую же последовательность аминокислотных остатков (гистоновый валик), что и гистоны. (Гистоны H3 и H4 образуют ядро гистонового октамера — главного комплекса, который связывает ДНК в эукариотическом хроматине; см. разд. 29.7 «Организация гистонового октамера».) Вместе с другими TAF, $TAF_{II}42$ и $TAF_{II}62$ могут создать основу для структуры, напоминающей гистоновый октамер; подобная структура может служить для неспецифичных к последовательности взаимодействий $TF_{II}D$ с ДНК. Гистоновый валик также используется в парных взаимодействиях между другими TAF.

Некоторые из TAF_{II} встречаются не только в $TF_{II}D$, но и в других комплексах. В частности, гистоноподобные TAF_{II} встречаются также в белковых комплексах, изменяющих структуру хроматина перед транскрипцией (см. разд. 30.7 «Гистонацетилазы ассоциированы с активаторами транскрипции»).

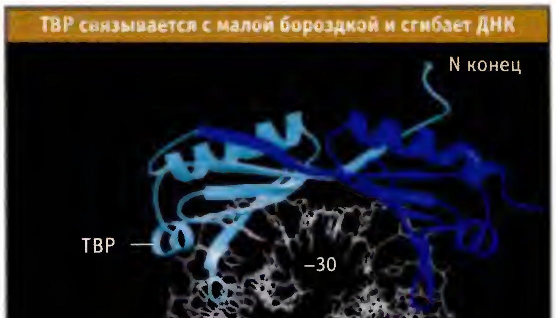


РИС. 24.13 Кристаллическая структура TBP с ДНК от –40 до стартовой точки показывает изгиб TATA-бокса, который расширяет малую бороздку, где связывается TBP. Фотография любезно предоставлена Stephen K. Burtley

24.10 Основной аппарат собирается на промоторе

Ключевые понятия:

- Первый шаг в инициации — связывание $TF_{II}D$ с ТАТА-боксом.
- Остальные факторы транскрипции связываются с комплексом в определенном порядке, увеличивая длину закрытого участка на ДНК.
- Когда РНК-полимераза II связывается с таким комплексом, она инициирует транскрипцию.

Для инициации необходимо, чтобы транскрипционные факторы действовали в определенном порядке и сформировали комплекс, который узнает РНК-полимераза. Порядок событий первоначально был определен путем наблюдения за увеличением размеров белкового комплекса, ассоциированного с ДНК. Сейчас возможно определять события более детально в терминах взаимодействий, которые видны на кристаллических структурах различных факторов, и в терминах связывания РНК-полимеразы с ДНК.

Путем футпринтинга районов ДНК, защищаемых каждым из комплексов, была получена модель, обобщенно представленная на РИС. 24.14. Как только какой-либо TF_{II}-фактор входит в комплекс, все более длинный участок ДНК оказывается закрытым. РНК-полимераза присоединяется на поздней стадии.

Сборка комплекса на промоторе начинается, когда $TF_{II}D$ связывается с ТАТА-боксом. ($TF_{II}D$ также распознает последовательность InR в стартовой точке.) Когда $TF_{II}A$ внедряется в комплекс, $TF_{II}D$ становится способен защитить вышележащий участок ДНК. $TF_{II}A$ может активировать ТВР путем ослабления репрессии, вызванной $TAF_{II}230$.

Добавление $TF_{II}B$ частично защищает кодирующую цепь вблизи стартовой точки, от –10 до +10 нуклеотида. Это позволяет предположить, что $TF_{II}B$ связывается ниже ТАТА-бокса, вероятно, асимметрично ориентированного относительно двух цепей ДНК. Кристаллическая структура, показанная на РИС. 24.15, позволяет расширить эту модель. $TF_{II}B$ связывается рядом с ТВР, увеличивая контакты вдоль одной стороны ДНК. Это создает контакты в малой бороздке выше ТАТА-бокса и контакты в большой бороздке выше ТАТА-бокса, в участке, называемом BRE. У архей гомолог $TF_{II}B$ действительно создает специфичные к последовательности контакты с промотором на участке BRE. Возможно, $TF_{II}B$ формирует поверхность, которая распознается РНК-полимеразой, таким образом, он отвечает за направление связывания фермента.

На кристаллической структуре $TF_{II}B$ с РНК-полимеразой видно, что три домена фактора взаимодействуют с ферментом. На РИС. 24.16 схематично представлено, как N-концевая цинковая лента $TF_{II}B$ соприкасается с ферментом возле места, где выходит РНК; возможно, это интерферирует с выходом РНК и влияет на переключение между абортивной инициацией и уходом с промотора. Удлиненный «палец» $TF_{II}B$

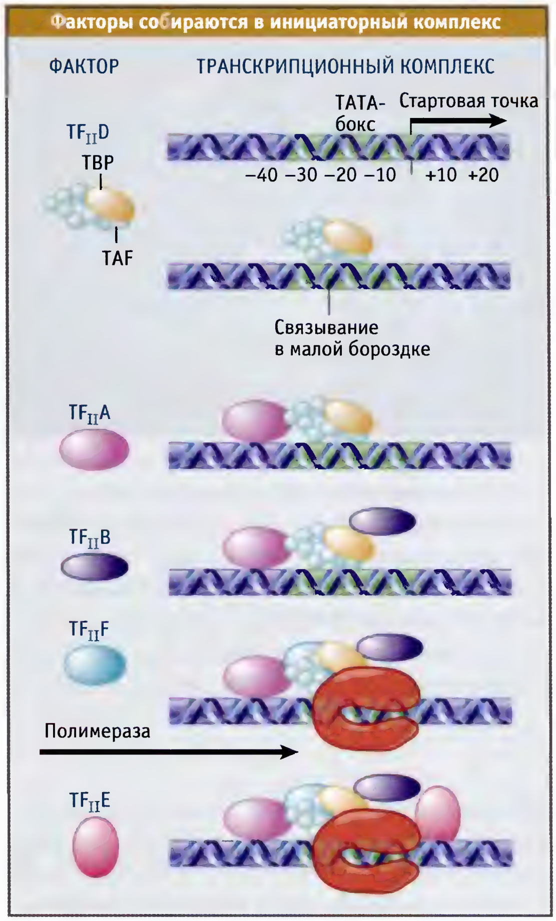
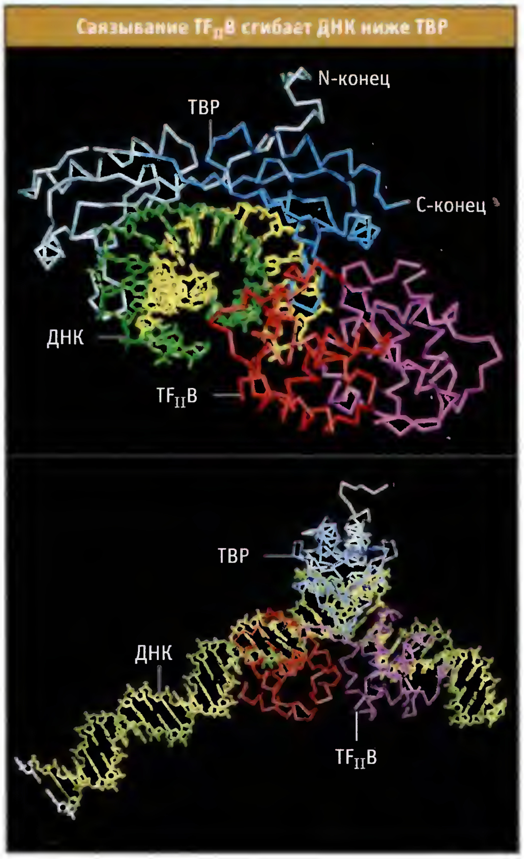


РИС. 24.14 **Инициаторный комплекс собирается на промоторах для РНК-полимеразы II путем последовательного организованного присоединения транскрипционных факторов**

внедряется в активный центр полимеразы. Для того чтобы правильно ориентировать ДНК, С-концевой домен взаимодействует с РНК-полимеразой и с $TF_{II}D$. Он также определяет фрагмент ДНК, контактирующий с факторами $TF_{II}E$, $TF_{II}F$, и $TF_{II}H$, и, по-видимому, располагает их на базовом комплексе факторов.

Фактор $TF_{II}F$ — гетеротетрамер, состоящий из двух типов субъединиц. Бóльшая субъединица (RAP74) имеет АТФ-зависимую ДНК-хеликазную активность, которая, возможно, используется в процессе плавления ДНК при инициации. Меньшая субъединица (RAP38) частично гомологична участкам бактериального сигма-фактора, контактирующим с ядром полимеразы; она плотно связывается с РНК-полимеразой II. Возможно, $TF_{II}F$ приносит РНК-полимеразу II к собирающемуся транскрипционному комплексу и обеспечивает возможность его объединения. Комплекс ТВР и TAF может взаимодействовать с CTD-хвостом РНК-полимеразы. Взаимодействие с $TF_{II}B$ также может оказаться важным, когда $TF_{II}F$/полимераза присоединяются к комплексу.

Связывание полимеразы удлиняет защищенный участок до +15 нуклеотида на кодирующей цепи и до +20 на некодирующей цепи.



На этих двух проекциях тройственного комлекса $TF_{II}B$-ТВР-ДНК показано, что $TF_{II}B$ связывается вдоль изогнутой поверхности ДНК. Две цепи ДНК показаны зеленым и желтым, ТВР показан синим, а $TF_{II}B$ красным и пурпурным. Фотографии любезно предоставлены Stephen K. Burtley

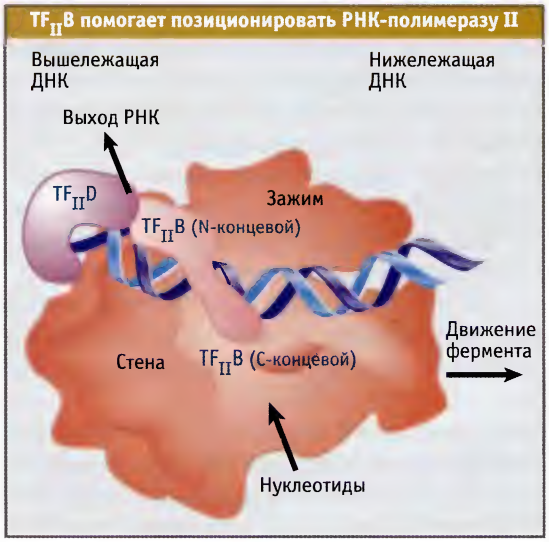
позволить основной полимеразе специфично распознать ДНК на промоторной последовательности, но при этом они развили бóльшую независимость. Конечно, факторы, в первую очередь, отвечают за специфичность при распознавании промотора. Лишь некоторые из факторов участвуют в контактах белок-ДНК (и только ТВР осуществляет сиквенс-специфичные контакты); таким образом, белок-белковые взаимодействия очень важны при сборке комплекса.

Когда ТАТА-бокс присутствует, он определяет расположение стартовой точки. Делеция ТАТА-бокса приводит к тому, что стартовая точка становится ошибочной, хотя общее снижение уровня транскрипции относительно невелико. Естественно, некоторые промоторы без ТАТА-бокса не имеют уникальных стартовых точек; вместо этого инициация происходит на любой точке из целого кластера стартовых точек. ТАТА-бокс позиционирует РНК-полимеразу (через взаимодействие с $TF_{II}D$ и другими факторами) таким образом, что она инициирует на правильном сайте. Это объясняет, почему его расположение фиксировано относительно стартовой точки. Связывание ТВР с ТАТА-боксом — главное событие при распознавании промотора, но два крупных TAF ($TF_{II}250$ и $TF_{II}150$) также контактируют с ДНК вблизи от стартовой точки и влияют на эффективность реакции.

Несмотря на то что *in vitro* сборка может происходить на ядре промотора, этой реакции недостаточно для транскрипции *in vivo*, где необходимы взаимодействия с активаторами, распознающими вышележащие элементы. Активаторы взаимодействуют с базовым аппаратом во время его сборки на разных стадиях (см. разд. 25.5 «Активаторы взаимодействуют с основным аппаратом транскрипции»).

Что же происходит на промоторах без ТАТА-бокса? Необходимы те же главные транскрипционные факторы, включая $TF_{II}D$. Участок Inr обеспечивает правильное расположение; с ним связывается $TF_{II}D$ благодаря способности одного или более TAF напрямую распознавать Inr. Другие TAF в $TF_{II}D$ также распознают DPE-элемент, расположенный ниже стартовой точки. Функция ТВР на этих промоторах сходна с его функцией на промоторах для РНК-полимеразы I и на внутренних промоторах для РНК-полимеразы III.

Сборка инициаторного комплекса РНК-полимеразы II являет собой контраст с транскрипцией у прокариотических организмов. Бактериальная РНК-полимераза — по сути, единый комплекс с изначально присущей ему способностью связываться с ДНК; сигма-фактор, необходимый для инициации, но не для элонгации, становится частью фермента до связывания с ДНК, хотя позднее и высвобождается. РНК-полимераза II может связываться с промотором, но только после того, как с ним свяжутся отдельные транскрипционные факторы. Эти факторы играют роль, аналогичную бактериальному сигма-фактору, — чтобы



$TF_{II}B$ связывается с ДНК и контактирует с РНК-полимеразой возле места выхода РНК и возле активного центра, и ориентирует ее на ДНК. Сравните с рис. 24.17, на котором показана структура РНК, участвующей в транскрипции

24.11 Инициация сопровождается освобождением промотора

Ключевые понятия:

- Для плавления ДНК и движения полимеразы необходимы $TF_{II}E$ и $TF_{II}H$.
- Для начала элонгации может потребоваться фосфорилирование CTD.
- Дальнейшее фосфорилирование CTD необходимо некоторым промоторам для завершения неудачной инициации.
- CTD может координировать процессинг РНК с транскрипцией.

Большинство главных транскрипционных факторов необходимы для связывания РНК-полимеразы с промотором, но некоторые действуют и на более поздней стадии. Связывание $TF_{II}E$ приводит к тому, что защищенная левая граница промоторной области сдвигается еще на один виток двойной спирали до +30. Два других фактора, $TF_{II}H$ и $TF_{II}J$, входят в комплекс после $TF_{II}E$. Они не изменяют паттерн связывания с ДНК.

$TF_{II}H$ — единственный главный транскрипционный фактор, обладающий множеством независимых ферментативных активностей. Его активности включают АТФазу, хеликазы обеих полярностей и киназную активность, которая может фосфорилировать CTD РНК-полимеразы II. $TF_{II}H$ — исключительный фактор еще и в том смысле, что он играет роль также и в процессе элонгации. Его взаимодействие с ДНК ниже стартовой точки необходимо для того, чтобы РНК-полимераза покинула промотор. Также $TF_{II}H$ участвует в репарации поврежденной ДНК (см. разд. 24.12 «Связь между транскрипцией и репарацией»).

Как было установлено по формированию первой фосфодиэфирной связи, реакция инициации происходит после связывания РНК-полимеразы. На рис. 24.17 предложена модель, согласно которой фосфорилирование С-конца РНК-полимеразы II необходимо для ее освобождения от транскрипционных факторов; таким образом, это может перевести процесс в стадию элонгации. Большинство транскрипционных факторов на этой стадии высвобождаются с промотора.

На линейной матрице для движения полимеразы необходимы: гидролиз АТФ, $TF_{II}E$ и хеликазная активность $TF_{II}H$ (благодаря субъединице XPB). Эти активности необязательны, если матрица сверхспирализована, а это предполагает, что $TF_{II}E$ и $TF_{II}H$ необходимы для плавления ДНК, чтобы полимераза могла начать движение. За непосредственное плавление ДНК отвечает хеликазная активность XPB-субъединицы $TF_{II}H$.

РНК-полимераза II приостанавливается после начала транскрипции некоторых генов. (Этот этап схож с абортивной инициацией бактериальной РНК-полимеразой, описанной в разд. 11.11 «Сигма-фактор контролирует связывание РНК-полимеразы с ДНК», хотя механизм различается.) После транскрипции небольшого участка на многих генах РНК-полимераза II терминирует. Короткий РНК-продукт быстро деградирует. Для продолжения элонгации всего гена необходимо действие киназы, называемой P-TEFb. Данная киназа является членом cdk-семейства, контролирующего клеточный цикл. P-TEFb действует на CTD для дальнейшего его фосфорилирования. Пока неясно, почему на одних промоторах это необходимо, а на других нет, и как этот процесс регулируется.

CTD также может участвовать, прямо или косвенно, в процессинге РНК после ее синтеза РНК-полимеразой II. На рис. 24.18 суммированы реакции процессинга, в которых CTD мог бы участвовать. Кэпирующий фермент (гуанилил-трансфераза), добавляющий остаток гуанидина к 5′-концу новосинтезированной мРНК, связывается с фосфорилированным CTD; это может быть важно потому, что позволяет ему модифицировать 5′-конец, как только тот синтезируется. Комплекс белков под названием SCAF связывается с CTD, а эти белки, в свою очередь, могут связываться с факторами сплайсинга. Это может быть средством координации транскрипции и сплайсинга. Некоторые компоненты аппарата расщепления/полиаденилирования

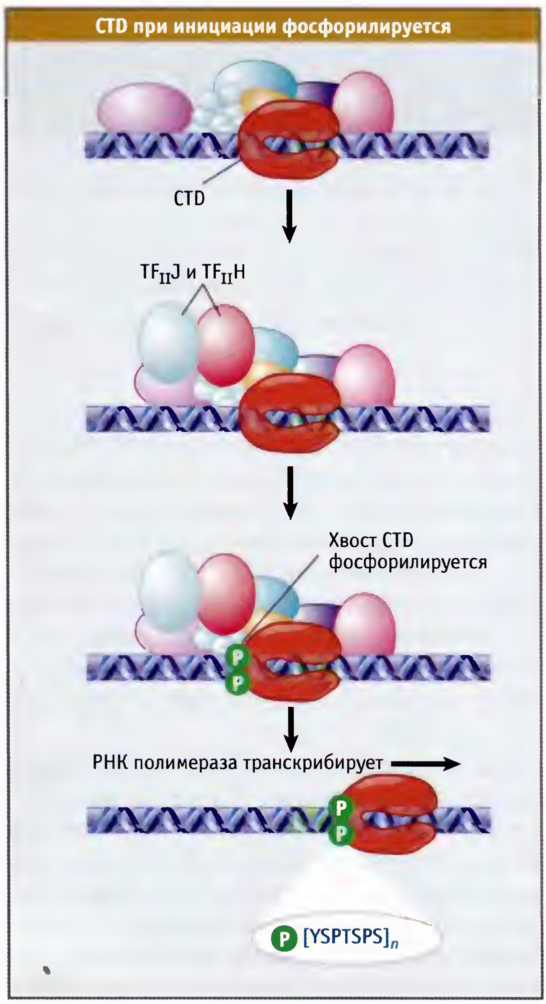


РИС. 24.17 Фосфорилирование CTD киназной активностью $TF_{II}H$, по-видимому, необходимо для высвобождения РНК-полимеразы и начала транскрипции

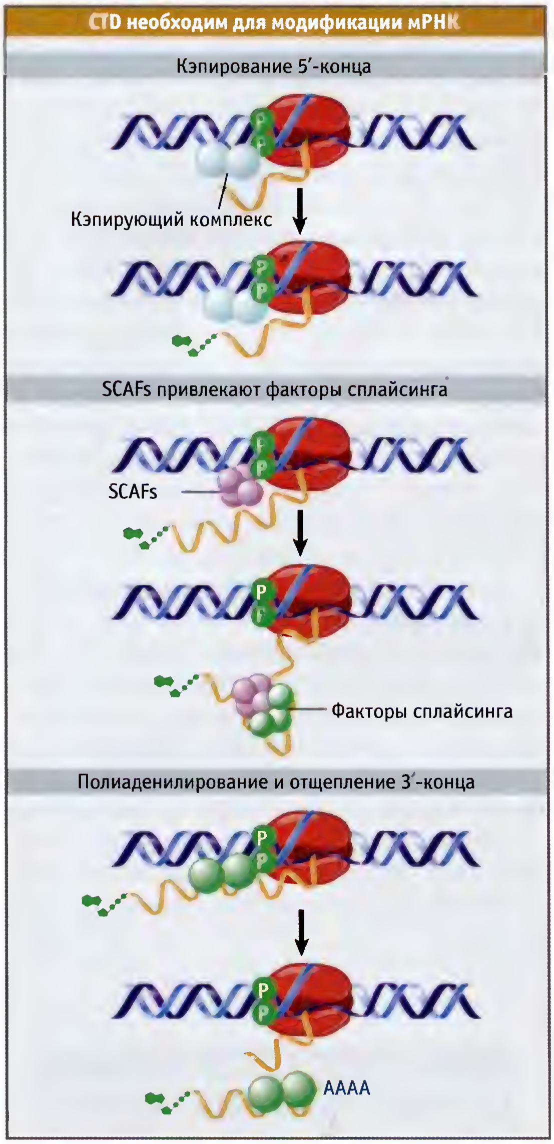


РИС. 24.18 **CTD очень важен для привлечения ферментов модификации РНК**

также связываются с CTD. Достаточно странно, что они это делают во время инициации, и, таким образом, РНК-полимераза готова к реакциям процессинга 3′-конца, как только она садится на матрицу! Все это говорит в пользу того, что CTD может быть центральным фактором, объединяющим процессинг с транскрипцией. В случае с кэпированием и сплайсингом CTD действует непрямым образом, вызывая образование белковых комплексов, которые и производят реакции. В случае с образованием 3′-конца он может участвовать в реакции напрямую.

В целом процесс инициации у эукариот сходен с процессом, катализируемым бактериальной РНК-полимеразой. Связывание РНК-полимеразы образует закрытый комплекс, который на более поздней стадии превращается в открытый комплекс, где цепи ДНК разделены. У бактерий образование открытого комплекса завершает необходимые структурные изменения в ДНК; специфика у эукариот заключается в том, что после формирования открытого комплекса необходимо дальнейшее раскручивание матрицы.

24.12 Связь между транскрипцией и репарацией

Ключевые понятия:

- Транскрибируемые гены в случае повреждения ДНК быстро репарируются.
- $TF_{II}H$ обеспечивает связь с комплексом ферментов репарации.
- Мутации компонента XPD в $TF_{II}H$ у человека вызывают три типа заболеваний.

И у бактерий, и у эукариот существует прямая связь между работой РНК-полимеразы и активацией репарации. Эта связь была обнаружена, поскольку чаще других репарируются активно транскрибируемые гены. Позже обнаружили, что только кодирующая цепь ДНК подвергается репарации, а некодирующая цепь репарируется с той же скоростью, что и основная масса ДНК.

У бактерий репарирующая активность обеспечивается системой вырезания-встраивания *uvr* (см. разд. 20.3 «Системы эксцизионной репарации *E. coli*»). Предпочтительная репарация блокируется мутациями в гене *mfd*, продукт которого обеспечивает связь РНК-полимеразы с ферментами Uvr.

На РИС. 24.19 представлена модель, показывающая связь между транскрипцией и репарацией. Когда РНК-полимераза встречает повреждения в ДНК, ее движение останавливается (поскольку она не может использовать поврежденную последовательность в качестве матрицы). Это объясняет, почему эффект специфичен к кодирующей цепи (повреждение в некодирующей цепи не препятствует действию РНК-полимеразы).

Белок Mfd выполняет две функции. Во-первых, он смещает тройственный комплекс РНК-полимеразы с ДНК. Во-вторых, он обеспечивает связывание фермента UvrABC с поврежденной ДНК. Это ведет к репарации ДНК с помощью механизма вырезания-встраивания (см. рис. 20.11). После того как ДНК подверглась репарации, следующая РНК-полимераза, транскрибирующая данный ген, образует уже нормальный транскрипт.

Подобный механизм, хотя и с использованием других ферментов, реализуется у эукариот. Предпочтительнее репарируется кодирующая цепь транскрибируемого гена, получившего повреждения из-за УФ-облучения. В этом процессе участвует $TF_{II}H$. Обнаружены альтернативные формы $TF_{II}H$, состоящие из ядра, связанного с другими субъединицами.

$TF_{II}H$ выполняет одну и ту же функцию и при инициации транскрипции, и при репарации повреждений

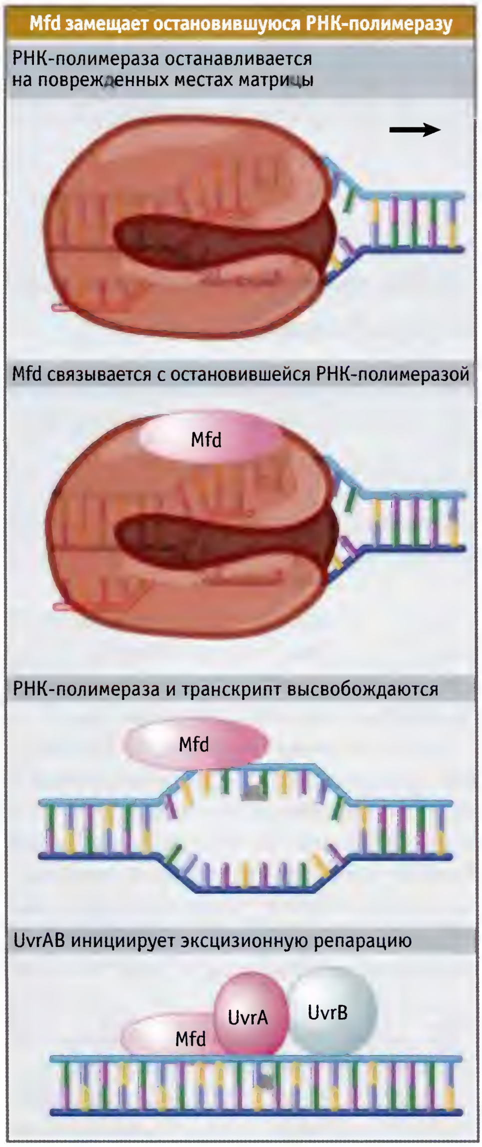


РИС. 24.19 **Mfd распознает остановившуюся РНК-полимеразу и направляет репарацию ДНК к поврежденной кодирующей цепи**

Одна и та же субъединица с хеликазной активностью (XPD) создает первый транскрипционный пузырек и расплавляет ДНК в поврежденном сайте. Другие его функции различаются в процессах транскрипции и репарации и обеспечиваются соответствующей формой комплекса.

На РИС. 24.20 показано, что главный фактор, участвующий в транскрипции, состоит из ядра (состоящего, в свою очередь, из пяти субъединиц), связанного с другими субъединицами, имеющими киназную активность; этот комплекс также включает в себя и субъединицу для репарации. Киназная каталитическая субъединица, фосфорилирующая CTD РНК-полимеразы, принадлежит к группе киназ, участвующих в контроле клеточного цикла. Возможно, эта связь влияет на транскрипцию и изменяет ее в соответствии с каждой конкретной стадией клеточного цикла.

Альтернативный комплекс состоит из ядра, связанного с большой группой белков, кодируемых генами репарации. (В общих чертах модель репарации показана на рис. 20.25.) Белки репарации включают особую субъединицу (XPC), распознающую поврежденную ДНК и обеспечивающую функцию присоединения, что позволяет предпочтительно репарировать кодирующую цепь, когда РНК-полимераза останавливается на поврежденных участках ДНК. Среди других белков, ассоциированных с комплексом, есть эндонуклеазы (XPG, XPF и ERCC1). Гомологичные белки найдены в комплексах у дрожжей (где их часто можно определить с помощью *rad*-мутаций, в которых нарушена репарация) и у человека (где их можно определить с помощью мутаций, вызывающих заболевания, являющиеся следствием дефектов репарации поврежденной ДНК). Субъединицы под названием XP кодируются генами, мутации в которых вызывают заболевание пигментную ксеродерму (см. разд. 20.11 «Системы репарации эукариотических клеток консервативны»).

Киназный комплекс и комплекс репарации могут обратимо ассоциировать и диссоциировать от ядра $TF_{II}H$. Это подразумевает модель, в которой одна форма $TF_{II}H$ необходима для инициации, но может быть замещена другой формой (возможно, при встрече с повреждением в ДНК). $TF_{II}H$ диссоциирует от РНК-полимеразы на ранней стадии элонгации (после транскрипции примерно 50 пн); для его повторного присоединения в месте повреждения ДНК могут потребоваться дополнительные компоненты.

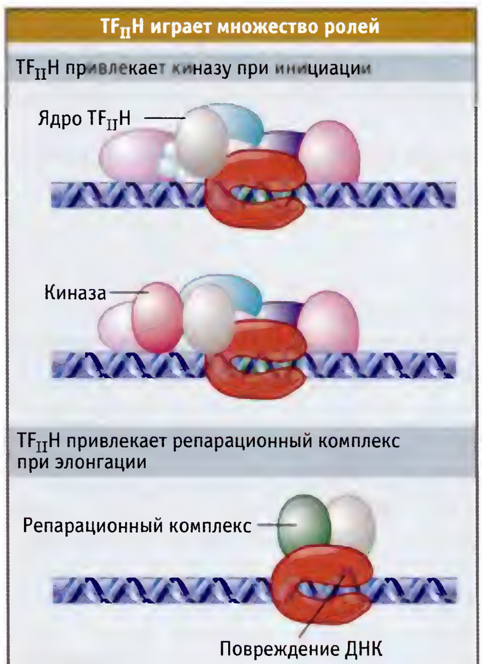


РИС. 24.20 **Ядро $TF_{II}H$ может соединяться с киназой при инициации и соединяться с репарационным комплексом при встрече с поврежденной ДНК**

Для обеспечения функции репарации может потребоваться модификация или даже деградация РНК-полимеразы. Большая субъединица РНК-полимеразы деградирует, когда фермент останавливается в местах УФ-повреждений. Мы пока не вполне понимаем связь между аппаратами транскрипции/репарации как таковыми и деградацией РНК-полимеразы. Возможно, удаление полимеразы необходимо, поскольку она остановилась.

Деградация РНК-полимеразы не происходит в клетках пациентов с синдромом Кокайна (расстройство репарации). Синдром Кокайна вызывается мутациями в одном из двух генов (*CSA* и *CSB*), продукты которых, по-видимому, являются частью или просто связаны с $TF_{II}H$. Синдром Кокайна также иногда вызывается мутациями в XPD.

Другое заболевание, которое может быть вызвано мутациями в XPD, — это трихотиодистрофия, которая имеет мало общего с пигментной ксеродермой или синдромом Кокайна (это заболевание влечет за собой умственную отсталость и характеризуется изменениями в структуре волос). Все это указывает на XPD как на плейотропный белок, разные мутации в котором могут влиять на разные функции. Фактически XPD необходим для стабильности комплекса $TF_{II}H$ во время транскрипции, однако хеликазная активность как таковая не является необходимой. Мутации, не позволяющие XPD стабилизировать комплекс, вызывают трихотиодистрофию. Для функции репарации требуется хеликазная активность. Мутации, влияющие на хеликазную активность, вызывают дефекты репарации, что приводит к пигментной ксеродерме или синдрому Кокайна.

24.13 Активаторы связываются с короткими элементами последовательности

Ключевые понятия:

- **Короткие консервативные элементы последовательности распределены на участках, предшествующих стартовой точке**
- **Вышележащие элементы увеличивают частоту инициации**
- **Факторы, которые связываются с ними для стимуляции транскрипции, называют активаторами.**

Промоторы для РНК-полимеразы II состоят из двух типов участков. Стартовую точку саму по себе определяют по расположению Inr и/или по близости ТАТА-бокса. Будучи связанной с главными транскрипционными факторами, РНК-полимераза II образует инициаторный комплекс, окружающий стартовую точку, как уже было описано. Однако эффективность и специфичность, с которыми распознается каждый конкретный промотор, зависят от коротких вышележащих последовательностей, которые, в свою очередь, распознаются специальной группой факторов, обычно называемых **активаторами**. Как правило, целевые последовательности лежат примерно в 100 пн выше стартовой точки, но иногда могут находиться и дальше. Связывание активаторов в этих сайтах может влиять на образование инициаторного комплекса на любой стадии.

Анализ типичного промотора показан на РИС. 24.21. Замены конкретных оснований проводились практически в каждой позиции на участке в 100 пн выше стартовой точки гена β-глобина. Уди-



РИС. 24.21 Насыщающий мутагенез вышележащего участка промотора β-глобина указывает на три коротких участка (в районе −30, −75 и -90), необходимых для инициации транскрипции. Эти участки соответствуют ТАТА-, CAAT- и GC-боксам

вительный результат этого эксперимента состоит в том, что *большинство мутаций не влияют на способность промотора инициировать транскрипцию*. Мутации, снижающие функцию промотора, расположены в трех положениях, соответствующих трем отдельным коротким элементам. Два вышележащих элемента оказывают больший эффект на уровень транскрипции, чем самый близкий к стартовой точке элемент. Мутации, усиливающие функцию, были найдены только в одном из элементов. Можно сделать вывод, что три коротких последовательности, в районе положений −30, −75 и −90, и составляют промотор. Каждая из них соответствует консенсусной последовательности для часто встречающихся типов промоторных элементов.

TATA-бокс (расположен на участке −30) — самый низкоэффективный компонент промотора, что было определено путем уменьшения транскрипции, вызываемой мутациями в нем. Обратите внимание, что хотя инициация и не прекращается при мутировании TATA-бокса, стартовая точка начинает варьировать и отклоняться от своего точного расположения. Это подтверждает роль TATA-бокса как основного позиционирующего компонента ядра промотора.

Базовые элементы и вышележащие элементы выполняют разные функции. Базовые элементы (TATA-бокс и Inr) главным образом определяют расположение стартовой точки. Они указывают *место*, где основные транскрипционные факторы собираются и образуют основной комплекс. Вышележащие элементы последовательности влияют на *частоту* инициации, действуя, вероятно, непосредственно на основные транскрипционные факторы, увеличивая эффективность их объединения в инициаторный комплекс (см. разд. 25.5 «Активаторы взаимодействуют с основным аппаратом транскрипции»).

Последовательность на участке −75 — это **CAAT-бокс** (названный по своей консенсусной последовательности). Это был один из первых изученных часто встречающихся элементов. Он обычно расположен вблизи −80, но может действовать и на других расстояниях от стартовой точки. CAAT-бокс выполняет свою функцию в любой ориентации. Чувствительность к мутациям указывает на то, что CAAT-бокс играет важную роль в определении эффективности промотора, но не влияет на его специфичность.

GC-бокс на участке −90 содержит последовательность GGGCGG. Часто в промоторе присутствуют множественные копии данного элемента, встречающиеся в разных ориентациях. GC-бокс также достаточно широко распространенный компонент промоторов.

24.14 Конструкция промотора может видоизменяться, но контекст имеет значение

Ключевые понятия:

- **Ни один из вышележащих элементов не является необходимым для функционирования промоторов, тем не менее хотя бы один элемент обязательно должен присутствовать для эффективной инициации.**
- **Некоторые элементы распознаются множеством факторов, и фактор, который используется каждым конкретным промотором, можно определить в контексте других, связывающихся с промотором факторов.**

Промоторы организованы по принципу «смешивания и соответствия». Разнообразные элементы могут влиять на функцию промотора, но ни один не является необходимым для всех промоторов. Некоторые примеры представлены на РИС. 24.22. Четыре типа элементов найдены в промоторах: TATA-, GC-боксы, CAAT-бокс и октамер (8-нуклеотидный элемент). Элементы, присутствующие в каждом конкретном промоторе, различны по числу, расположению и ориентации. Ни один элемент не является общим для всех промоторов. Несмотря на то что промоторы передают направленную информацию (ведь транскрипция осуществляется только в одном направлении — вниз от стартовой точки), GC- и CAAT-боксы, по-видимому, способны выполнять свои функции в обеих ориентациях. Это указывает, что элементы функционируют сами по себе как ДНК-связывающие сайты, чтобы приблизить транскрипционные факторы к стартовой точке; структура фактора должна быть достаточно гибкой, чтобы позволять осуществлять белок-белковые контакты с основным аппаратом независимо от ориентации его ДНК-связывающего домена и расстояния от стартовой точки.

Более или менее широко распространенные активаторы доступны для любого промотора, который имеет копию распознаваемого ими элемента. Общие элементы,

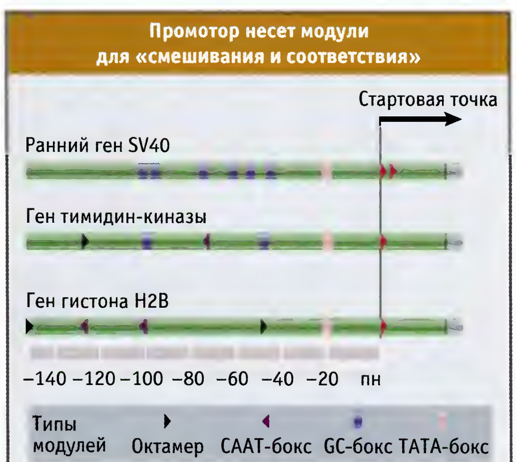


РИС. 24.22 Промоторы имеют различные комбинации TATA-боксов, CAAT-боксов, GC-боксов и других элементов

распознаваемые часто встречающимися активаторами, включают: CAAT-бокс, GC- бокс и октамер. Вероятно, всем промоторам требуется один или более этих элементов, чтобы эффективно функционировать. Активатор обычно связывает консенсусную последовательность размером <10 пн, но в действительности покрывает участок ДНК длиной около 20 пн. Зная размеры активаторов и длину участка ДНК, который каждый из них покрывает, можно ожидать, что различные белки все вместе покроют весь участок выше стартовой точки (на котором расположены консервативные элементы).

В целом определенная консенсусная последовательность распознается соответствующим активатором (или членом семейства факторов). Однако иногда случается, что определенная промоторная последовательность распознается одним из нескольких активаторов. Широко распространенный активатор, Oct-1, связывается с октамером для активации гена гистона H2B (и предположительно, ряда других генов). Oct-1 — это единственный фактор, связывающий октамер в нелимфоидных клетках. В лимфоидных клетках существует другой активатор, Oct-2, который связывается с октамером для активации гена легкой цепи иммуноглобулина κ. Таким образом, Oct-2 представляет собой тканеспецифичный активатор, в то время как Oct-1 — универсальный. Следовательно, различные активаторы распознают CAAT-бокс.

Использование одного и того же октамера для экспрессирующегося во всех клетках гена H2B и лимфоид-специфичного гена иммуноглобулина представляет собой парадокс. Почему вездесущий Oct-1 неспособен активировать гены иммуноглобулинов в нелимфоидных тканях? Здесь важен *контекст*: Oct-2, а не Oct-1, может быть необходим для взаимодействия с другими белками, связывающимися с промотором. Этот пример показывает, что мы не можем предсказать, будет ли ген активироваться определенным активатором только на основе присутствия конкретных элементов последовательности в его промоторе.

24.15 Энхансеры несут двунаправленные элементы, которые способствуют инициации

Ключевые понятия:

- Энхансер активирует ближайший к нему промотор, расположенный на любом расстоянии выше или ниже него.
- UAS (upstream activator sequence, вышележащая активаторная последовательность) у дрожжей похожа на энхансер, но работает только выше промотора.
- В энхансерах и промоторах обнаруживаются схожие элементы последовательности.
- Энхансеры образуют комплексы с активаторами, которые прямым или косвенным образом взаимодействуют с промоторами.

До сих пор мы рассматривали промотор как изолированный участок, отвечающий за связывание с РНК-полимеразой. Однако эукариотические промоторы не обязательно функционируют в одиночку. Как минимум в некоторых случаях активность промотора чрезвычайно усиливается в присутствии энхансера, состоящего из отдельной группы элементов, но расположенного на различном расстоянии от самого промотора.

Представление о том, что энхансер отличается от промотора, отражает две характеристики. Позиция энхансера относительно промотора может существенно различаться. На рис. 24.23 показано, что энхансер может быть как выше, так и ниже промотора. Кроме того, он может выполнять свои функции в любой ориентации (то есть он может быть инвертирован) относительно промотора. Опыты с ДНК показали, что энхансер может стимулировать работу любого промотора, помещенного неподалеку. В естественных условиях в геноме энхансеры могут быть расположены внутри генов (то есть ниже

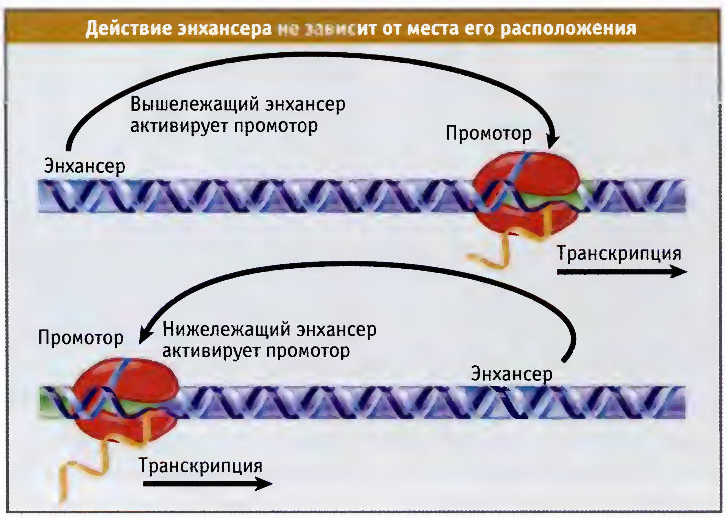


Рис. 24.23 Энхансер может активировать промотор из вышележащего или нижележащего положения, и его последовательность может быть инвертирована относительно промотора

промотора) или на расстоянии в десятки тысяч пн в любом направлении.

Для рабочих целей иногда полезно определить промотор как *последовательность или последовательности ДНК, которые должны находиться в (относительно) фиксированном положении относительно стартовой точки*. По этому определению TATA-бокс и другие вышележащие элементы включаются в промотор, а энхансер исключается. Однако это только рабочее определение, а не точная классификация.

У дрожжей обнаружены элементы, аналогичные энхансерам, которые назвали **вышележащими активаторными последовательностями** (**UAS**, upstream activator sequences). Они могут функционировать в любой ориентации на различных расстояниях выше промотора, но не могут выполнять свои функции, когда расположены ниже. Они играют регуляторную роль: в некоторых случаях UAS связываются с регуляторным белком (или белками), которые активируют нижележащие гены.

Опыты по перестановке элементов в ДНК показали, что если вырезать энхансерную последовательность из ДНК и потом поместить ее в другое место, то нормальная транскрипция будет поддерживаться так долго, пока эта последовательность присутствует *где угодно* в молекуле ДНК. Если ген β-глобина поместить в молекулу ДНК, несущую энхансер, *in vivo* его транскрипция усиливается более чем в 200 раз, даже когда энхансер расположен в нескольких тпн выше или ниже стартовой точки в любой ориентации. На какой дистанции энхансер перестает работать — пока неясно.

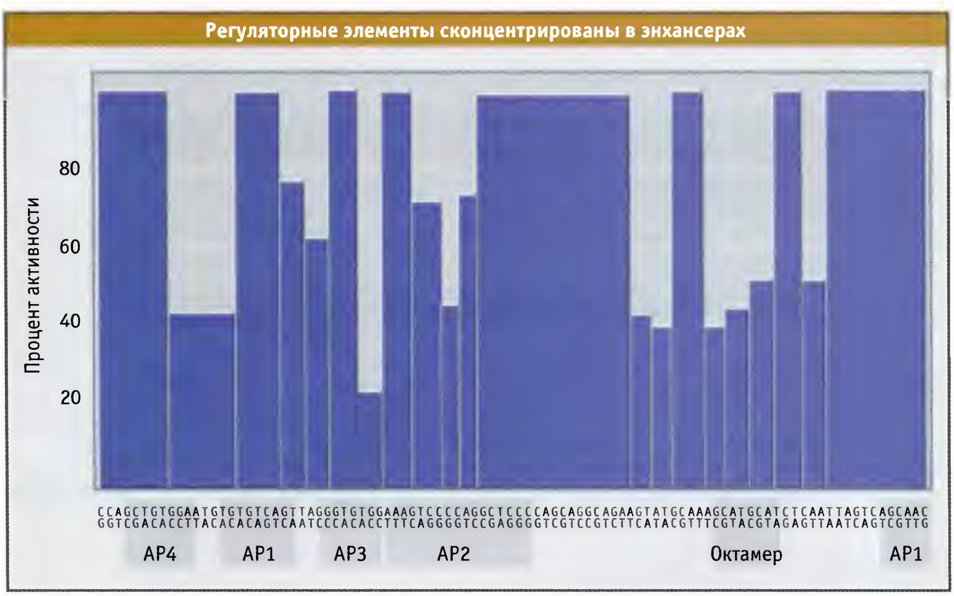
24.16 Энхансеры содержат те же элементы, что и промоторы

Ключевые понятия:

- **Энхансеры содержат те же самые короткие элементы последовательности, которые обнаружены в промоторах**
- **Плотность элементов последовательности в энхансерах выше, чем в промоторах.**

Разница между энхансером и типичным промотором проявляется в разнице в плотности регуляторных элементов. На показана чувствительность энхансера SV40 к повреждению мутагенезом. Можно видеть, что более высокая пропорция сайтов напрямую влияет на его функцию, чем в случае с промотором, проанализированным тем же способом на рис. 24.21. Присутствует увеличение плотности белок-связывающих сайтов. Многие из этих сайтов — обычные элементы промоторов; к примеру, AP1 или октамер.

Специфичность транскрипции может контролироваться как промотором, так и энхансером. Промотор может специфично регулироваться, а ближайший энхансер увеличивать эффективность инициации, либо у промотора может не быть специфичной регуляции, но он может становиться активным, только когда ближайший энхансер специфично активируется. Для примера возьмем гены иммуноглобулинов, у которых энхансер расположен *внутри* транскрипционной единицы. Энхансеры иммуноглобулинов, по-видимому, активны только в В-лимфоцитах, в которых работают гены иммуноглобулинов. Эти энхансеры обеспечивают часть регуляторной сети, с помощью которой контролируется экспрессия генов.



Энхансер содержит несколько структурных мотивов. На гистограмме отражен эффект мутаций, снижающих функцию энхансера до <75% от дикого типа. Сайты посадки белков показаны внизу гистограммы

Различие между промотором и энхансером состоит еще и в том, что энхансер обнаруживает бо́льшую кооперативность между связывающими факторами. Комплекс на энхансере, отвечающем на интерферон γ (ИФНγ), собирается кооперативно и образует функциональную структуру под названием **энхансеросома**. Связывание негистонового белка HMGI(Y) изгибает ДНК, после чего эта структура связывает несколько активаторов (NF-κB, IRF и ATF-Jun). В отличие от конструкции промотора по типу «смешивания и соответствия» все эти компоненты необходимы для создания активной структуры на энхансере. Сами по себе компоненты энхансера не связываются напрямую с РНК-полимеразой, но они создают поверхность, которая связывает *коактивирующий комплекс*. Этот комплекс помогает преинициаторному комплексу главных транскрипционных факторов, формирующемуся на промоторе, рекрутировать РНК-полимеразу. Более детально функцию коактиваторов мы обсудим в разд. 25.5 «Активаторы взаимодействуют с основным аппаратом транскрипции».

24.17 Функция энхансеров состоит в увеличении концентрации активаторов возле промотора

Ключевые понятия:

- Обычно энхансеры функционируют только в *цис*-положении относительно промотора-мишени.
- Можно заставить энхансеры работать и в *транс*-положении, соединив ДНК, несущую целевой промотор, с ДНК, несущую энхансер, через белковый мостик, или если эти две кольцевые молекулы будут сцеплены друг с другом.
- Суть в том, что энхансер будет функционировать в любом случае, если он находится недалеко от промотора.

Каким же образом энхансер способен стимулировать промотор, который может находиться на любом расстоянии от него и с любой стороны? Когда энхансеры были впервые открыты, рассматривались несколько возможностей для их действия как элементов, четко отличающихся от промотора:

- Возможно, энхансер изменяет общую структуру матрицы — к примеру, путем влияния на степень сверхспирализации.
- Возможно, энхансер отвечает за расположение матрицы в специальном месте в клетке — к примеру, присоединяя ее к ядерному матриксу.
- Может быть, энхансер обеспечивает «точку входа», то есть такое место, в котором РНК-полимераза (или какие-то другие необходимые белки) присоединяется к хроматину для инициации.

Сейчас мы считаем, что функция энхансера использует тот же тип взаимодействия с основным аппаратом, что и взаимодействие вышележащего промоторного элемента (UPE). Энхансеры представляют собой модульную структуру, как и промоторы. Некоторые элементы встречаются как в энхансерах, так и в промоторах. Некоторые отдельные элементы, обнаруживаемые в промоторах, разделяют с энхансерами способность функционировать на различном расстоянии и в любой ориентации. Таким образом, разница между энхансером и промотором размывается: энхансеры можно рассматривать как структуры, содержащие промоторные элементы, близко сгруппированные друг с другом, и способные функционировать на значительных расстояниях от стартовой точки.

Важная роль энхансеров может состоять еще и в том, что они увеличивают концентрацию активаторов вблизи промотора («близость к промотору» в данном случае — термин относительный). Два эксперимента, схемы которых представлены на РИС. 24.25, указывают, что это действительно так.

Фрагмент ДНК, содержащий энхансер на одном своем конце и промотор на другом конце, транскрибируется неэффективно; но энхансер может стимулировать транскрипцию с этого промотора, если их соединить белковым мостиком. Структурные эффекты, такие как изменение сверхспирализации, через такой мостик передать невозможно; и это доказывает, что в данном случае критичным моментом оказывается именно физическое сближение энхансера и промотора.

Энхансер у бактерий является сайтом связывания для регуляторного белка NtrC, который действует на РНК-полимеразу, использующую промоторы, распознаваемые фактором σ^{54}. Если энхансер поместить на кольцевую ДНК, сцепленную (образующую катенан) с кольцевой ДНК, несущей промотор, то инициация почти так же эффективна, как и в случае, когда энхансер и промотор расположены на одной и той же кольцевой молекуле. Однако в таком эксперименте не наблюдается никакой инициации транскрипции, если энхансер и промотор расположены на отдельных, не сцепленных между собой кольцевых молекулах. Это снова подтверждает, что критичным моментом является такое расположение белка, связанного с энхансером, которое увеличивает шансы энхансера на контакт с белком, связанным с промотором.

Если белки, связанные с энхансером, расположенным в нескольких тпн от промотора, напрямую взаимодействуют с белками, связанными вблизи стартовой точки, ДНК должна быть достаточно гибко организована, чтобы позволить энхансеру и промотору расположиться близко друг от друга. Для этого ДНК должна образовать большую петлю. И такие петли, действительно, можно непосредственно наблюдать в случае бактериального энхансера.

Существует интересное исключение из правила о действии энхансеров только в *цис*-положении в естественных условиях. Спаривание соматических хромосом позволяет энхансеру на одной хромосоме активировать промотор на партнерской хромосоме. И это подтверждает теорию о работе энхансеров только в непосредственной близости к промотору.

Что же ограничивает активность энхансера? Обычно он обеспечивает активацию ближайшего промотора. Но есть ситуации, когда энхансер располагается между

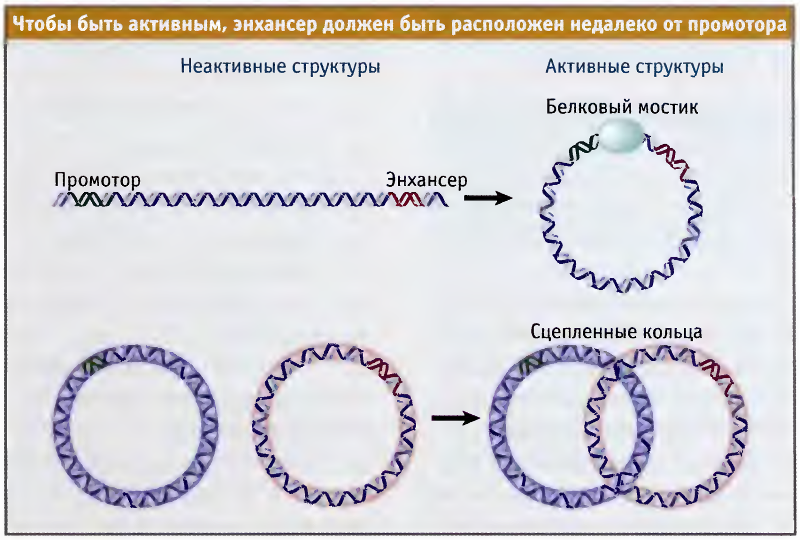


РИС. 24.25 Энхансер может приближать транскрипционные факторы к промотору и, таким образом, выполнять свои функции. Энхансер не действует на промотор с противоположного конца длинной линейной ДНК, но становится активен, когда ДНК образует кольцо с помощью белкового мостика. Энхансер и промотор на отдельных кольцевых ДНК не взаимодействуют, но могут начать взаимодействовать в случае, когда эти две молекулы сцеплены

двумя промоторами и активирует только один из них на основе специфичных белок-белковых взаимодействий между комплексами, сформированными на этих двух элементах. Активность энхансера может быть ограничена инсулятором — элементом ДНК, блокирующим действие энхансера на расположенный за ним промотор (см. разд. 29.14 «Инсуляторы блокируют действие энхансеров и гетерохроматина»).

Как именно усиливается процесс инициации с помощью энхансеров — пока не вполне ясно. Неизвестно, какая часть клеточных промоторов нуждается в энхансерах, чтобы иметь обычный для них уровень экспрессии генов, а также как часто энхансер является регуляторным элементом. Некоторые энхансеры активируются только в тканях, в которых функционируют их целевые гены, а другие могут быть активны во всех клетках.

24.18 Экспрессия генов связана с деметилированием

Ключевые понятия:

- Деметилирование 5′-конца гена необходимо для транскрипции

Метилирование ДНК — одно из нескольких регуляторных явлений, влияющих на активность промотора. Метилирование промотора блокирует транскрипцию, и для активации промотора метильные группы должны быть удалены. Этот эффект описан как для промоторов РНК-полимеразы I, так и для промоторов РНК-полимеразы II. Метилирование — обратимое регуляторное явление. Оно начинается с модификации гистонов, что включает в себя деацетилирование и метилирование белков (см. разд. 30.9 «Метилирование гистонов и метилирование ДНК взаимосвязаны»).

Метилирование бывает также и эпигенетическим. В этом случае модификация может специфично производиться в сперматозоидах или в ооцитах, и в результате в следующем поколении могут сохраняться различия между двумя аллелями. Это может приводить к различиям в экспрессии отцовского и материнского аллелей (см. разд. 31.8 «Метилирование ДНК обусловливает импринтинг»).

В этой главе мы рассмотрим способы, благодаря которым метилирование влияет на транскрипцию и которые являются универсальными, будь метильная группа добавлена или удалена в качестве локального регуляторного события или эпигенетического.

Метилирование оснований на промоторах для РНК-полимеразы II происходит на CG-динуклеотидах. Распределение метильных групп можно исследовать путем расщепления ДНК с помощью ферментов рестрикции, узнающих сайты, несущие CG-динуклеотиды. На РИС. 24.26 сравниваются два типа активностей рестриктаз. Такие ферменты называют **изошизомерами** — они расщепляют одну и ту же целевую последовательность в ДНК, но по-разному отвечают на ее метилирование.

Фермент HpaII расщепляет последовательность CCGG (указана последовательность только одной цепи ДНК). Однако если второй C метилирован, то фермент не может распознать этот сайт. А фермент MspI расщеп-

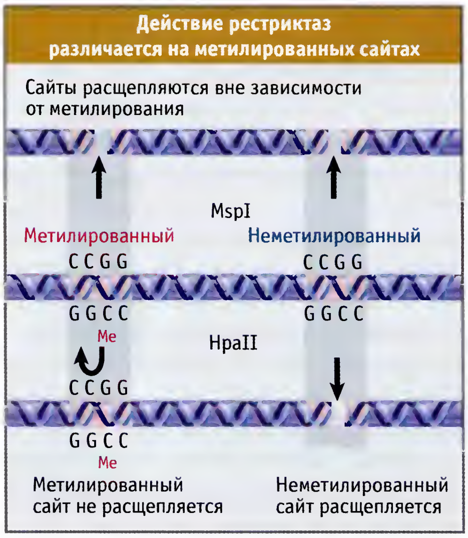


РИС. 24.26 **Фермент рестрикции MspI расщепляет все последовательности CCGG вне зависимости от их метилирования по второму C, а HpaII расщепляет только неметилированный тетрамер CCGG**

ляет тот же самый сайт вне зависимости от метилирования по C. Таким образом, MspI можно использовать для нахождения всех последовательностей CCGG, а HpaII можно использовать для определения, метилированы они или нет.

Используя неметилированную ДНК в качестве субстрата, эти два фермента образуют одинаковые фрагменты после рестрикции. Однако на метилированной ДНК модифицированные сайты не будут расщепляться HpaII. Для каждого такого сайта один большой фрагмент после обработки HpaII заменит два фрагмента от обработки MspI. На РИС. 24.27 приведен пример такого эксперимента.

Для многих генов состояние их метилирования постоянно на большинстве сайтов, но в ряде случаев варьирует. Некоторые сайты метилированы во всех исследованных тканях; некоторые сайты неметилированы во всех тканях. *Небольшое количество сайтов метилировано в тканях, в которых ген не экспрессируется, но не метилировано в тканях, в которых ген активен.* Таким образом, активный ген можно описать как *частично метилированный.*

Косвенное доказательство того, что деметилирование может привести к экспрессии генов, дали эксперименты с лекарством 5-азацитидином. Это вещество внедряется в ДНК вместо цитидина и не может быть метилировано, поскольку его 5′-положение блокировано. Это приводит к появлению деметилированных сайтов в ДНК после репликации (следуя правой схеме на рис. 15.7).

Фенотипические эффекты 5-азацитидина включают индуцирование изменений в клеточной дифференцировке. К примеру, индуцируется развитие мышечных клеток из немышечных клеток-предшественников. Вещество также активирует гены на молчащей Х-хромосоме; это позволяет предположить, что состояние метилирования может быть связано с неактивностью хромосомы.

Кроме исследования статуса метилирования резидентных генов можно сравнить результаты при внедрении метилированной и неметилированной ДНК в хозяйские клетки. Такие эксперименты показывают четкую корреляцию: *метилированные гены неактивны, а неметилированные гены активны.*

Какова степень метилирования на участке при частичном метилировании? У кур в генном кластере α-глобина в зрелых эритроидных клетках частичное метилирование ограничено участком от ~500 пн выше первого из двух зрелых α-генов до ~500 пн ниже второго. Сайты частичного метилирования присутствуют на всем участке, включая спейсер между генами. Участок частичного метилирования совпадает с участком максимальной чувствительности матрицы к ДНКазе I. Этот факт ставит под сомнение гипотезу о том, что частичное метилирование характерно для участка, который содержит транскрибируемый ген или гены. Так же как и другие изменения хроматина, кажется вероятным, что отсутствие метильных групп связано скорее со *способностью к транскрипции*, чем с самим актом транскрипции.

Проблема в понимании общей ассоциации между частичным метилированием и активацией генов в том, что используется лишь небольшая часть (иногда совсем небольшая часть) метилированных сайтов. Возможно, метилирование важно только на определенных сайтах или на определенном ограниченном участке. Возможно также, что снижение уровня метилирования (или даже полное удаление метильных групп с какого-то участка ДНК) — это часть неких структурных изменений, необходимых для осуществления транскрипции.

В частности, деметилирование промотора может быть необходимо для того, чтобы сделать его доступным для инициации транскрипции. В гене γ-глобина, к примеру, присутствие метильных групп на участке вокруг стартовой точки, между нуклеотидами −200 и +90, супрессирует транскрипцию. Удаление трех



РИС. 24.27 **Сравнение результатов расщепления ДНК ферментами MspI и HpaII на электрофорезе в агарозном геле**

метильных групп, расположенных выше стартовой точки, или трех метильных групп, расположенных ниже, не снимает супрессию. Однако удаление всех метильных групп позволяет промотору функционировать нормально. Таким образом, для транскрипции может быть необходим свободный от метильных групп промоторный участок (см. разд. 24.19 «CpG-островки являются мишенями для регуляции»). Но существуют и исключения из этого правила.

Некоторые гены могут экспрессироваться, даже когда они сильно метилированы. Таким образом, связь между метилированием и экспрессией в организме не универсальна, но главное правило состоит в том, что метилирование не позволяет генам экспрессироваться, а для экспрессии необходимо деметилирование.

24.19 CpG-островки являются мишенями для регуляции

Ключевые понятия:

- **Неметилированные CpG-островки окружают промоторы конститутивно экспрессирующихся генов.**
- **CpG-островки обнаруживаются также в промоторах некоторых тканеспецифичных генов.**
- **В геноме человека примерно 29 000 CpG-островков.**
- **Метилирование CpG-островка не позволяет промотору внутри него активироваться.**
- **Репрессия вызывается белками, связывающимися с CpG-динуклеотидами.**

Присутствие **CpG-островков** на 5′-концах некоторых генов связано с влиянием метилирования на экспрессию генов. Эти островки определяют по наличию увеличенной плотности динуклеотидной последовательности, CpG (CpG = 5′-CG-3′).

CpG-динуклеотиды присутствуют в ДНК позвоночных только в ~20% от той частоты, которую следует ожидать на основании расчета частоты встречаемости G-C последовательности. (Этот эффект может быть следствием того, что CpG-динуклеотиды метилированы по C, а спонтанное дезаминирование метил-C превращает его в T, что вносит мутацию и приводит к исчезновению динуклеотида.) Однако в некоторых участках плотность CpG-динуклеотидов достигает расчетного значения; фактически она там увеличена на порядок относительно остального генома. CpG-динуклеотиды на этих участках в основном не метилированы.

Такие CpG-богатые островки имеют содержание G-C сочетаний ~60%, по сравнению с 40% в среднем по всему геному. Они имеют форму участков ДНК длиной обычно от 1 до 2 тпн. В человеческом геноме ~45 000 таких островков. Некоторые из островков присутствуют в повторяющихся Alu-элементах и могут быть следствием высокого содержания G-C-сочетаний. За исключением этих случаев, в человеческом геноме остается ~29 000 островков. В мышином геноме их меньше, ~15 500. Около 10 000 предсказанных островков в обоих видах расположены, по-видимому, в контексте консервативных последовательностей; этот факт говорит о том, что такие островки могут играть важную регуляторную роль. Структура хроматина на этих участках несет признаки, характерные для экспрессирующихся районов (см. разд. 30.11 «Активации промотора представляет собой упорядоченную серию событий»); там понижен уровень гистона H1 (это, вероятно, означает, что структура менее компактизована), другие гистоны сильно ацетилированы (особенность, которая ассоциирована с экспрессией генов), а также там присутствуют гиперчувствительные сайты (как можно было бы ожидать для активных промоторов).

В некоторых случаях CpG-богатые островки начинаются чуть выше промотора и продолжаются ниже в транскрибируемый район, пока не исчезнут. На РИС. 24.28 сравнивается плотность CpG-динуклеотидов в «обычном» участке генома с CpG-островками, определенными по последовательности ДНК. CpG-островки окружают 5′-район гена APRT, который конститутивно экспрессируется.

Все конститутивно экспрессируемые гены «домашнего хозяйства» содержат CpG-островки; здесь находится около половины островков. Другая половина островков встречается в промоторах тканеспецифичных генов; меньшая часть (<40%) этих генов имеет островки. В этих случаях островки не метилированы вне зависимости от состояния экспрессии гена. Таким образом, присутствие неметилированных CpG-островков можно считать индикатором того, что ген

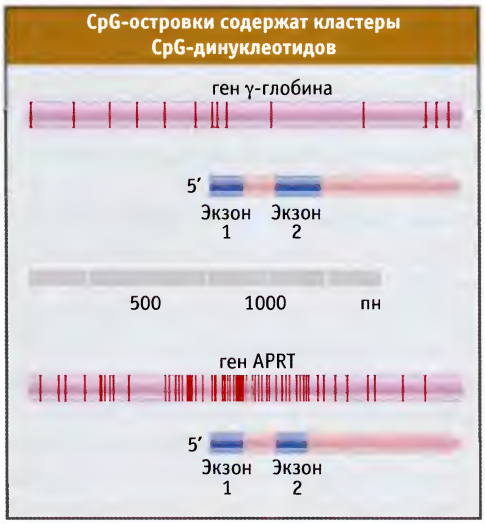


РИС. 24.28 Обычная плотность дублетов CpG в ДНК млекопитающих примерно 1/100 пн, как можно видеть на примере гена γ-глобина. На CpG-богатом островке плотность увеличивается до >10 дублетов/100 пн. Островок в гене APRT начинается примерно на 100 пн выше промотора и продолжается в гене примерно на 400 пн. Каждая вертикальная линия показывает положение CpG-дублета

является потенциально активным, но не обязательно транскрибируемым. Многие островки, не метилированные в организме животного, становятся метилированными в культуре клеток, и это может быть связано с неспособностью этих клеточных линий выполнять все функции, типичные для ткани, из которой они ведут свое происхождение.

Метилирование CpG-островков может влиять на транскрипцию. При этом может использоваться один из двух механизмов:

- Метилирование сайта связывания какого-либо фактора может запретить его связывание с ДНК. Это происходит в случае связывания с регуляторным сайтом, отличным от промотора (см. разд. 31.9 «Импринтинг генов с взаимно-противоположной половой специфичностью может контролироваться из одного центра на хромосоме»).
- Метилирование может вызвать связывание специфических репрессоров с ДНК.

Репрессия вызывается любым из двух типов белков, связывающихся с метилированными CpG-последовательностями. Для связывания белка MeCP1 с ДНК необходимо присутствие нескольких метильных групп, а белок MeCP2 и семейство родственных белков могут связываться с единственной метилированной парой CpG. Этим объясняется, почему для инициации транскрипции необходим свободный от метилирования участок. Связывание белков любого типа останавливает транскрипцию в ядерном экстракте *in vitro*.

MeCP2, репрессирующий транскрипцию путем взаимодействия с комплексами на промоторе, связан с репрессорным комплексом Sin3, имеющим гистон деацетилазную активность (см. рис. 30.16). Это наблюдение демонстрирует прямую связь между двумя типами репрессирующих модификаций: метилированием ДНК и деацетилированием гистонов.

Отсутствие метильных групп связано с экспрессией генов. Однако есть некоторые трудности в доказательстве того, что состояние метилирования является ключевым способом контроля экспрессии. У *Drosophila melanogaster* (и других видов отряда Двукрылые) ДНК очень слабо метилирована (хотя есть ген, вероятно, кодирующий метилтрасферазу), а у нематоды *Caenorabditis elegans* ДНК вообще не метилирована. Остальные различия между неактивным и активным хроматином у них, по-видимому, такие же, как и у видов, метилирующих ДНК. Таким образом, у этих организмов любая роль, какую бы метилирование ни играло у позвоночных, заменена какими-то другими механизмами.

Таким образом, в активных генах происходят три изменения:

- Возле промотора имеется гиперчувствительный сайт (сайты).
- Нуклеосомы в домене, включающем транскрибируемый участок, становятся более чувствительными к ДНКазе I.
- ДНК этого же участка частично метилирована.

Для транскрипции необходимы все эти изменения.

24.20 Резюме

Из трех РНК-полимераз эукариотических организмов, РНК-полимераза I транскрибирует рРНК и обеспечивает основную часть синтетической активности, РНК-полимераза II транскрибирует структурные гены с образованием мРНК, и РНК-полимераза III транскрибирует малые РНК. Эти ферменты имеют схожие структуры из двух больших субъединиц и множества более мелких субъединиц; есть и некоторые общие субъединицы.

Ни одна из трех РНК-полимераз напрямую не распознает свои промоторы. Общий принцип состоит в том, что факторы транскрипции несут главную ответственность за распознавание характерных элементов последовательности каждого конкретного промотора. По сути они служат для привлечения РНК-полимеразы и для правильного ее позиционирования на стартовой точке. На каждом типе промоторов инициаторный комплекс собирается в серии реакций, в которых отдельные факторы присоединяются (или покидают) комплекс. Фактор TBP необходим для инициации всех трех РНК-полимераз. В каждом случае он обеспечивает одну субъединицу транскрипционного фактора, который связывается вблизи от стартовой точки.

Промотор состоит из ряда коротких элементов последовательности на участке выше стартовой точки. Каждый элемент связывается транскрипционным фактором. Основной аппарат, состоящий из факторов TF_{II}, собирается на стартовой точке и позволяет связаться РНК-полимеразе. TATA-бокс возле стартовой точки (если он есть) и инициаторный участок непосредственно на стартовой точке отвечают за выбор конкретной стартовой точки на промоторах РНК-полимеразы II. TBP связывается напрямую с TATA-боксом (если он есть); на промоторах без TATA-бокса он располагается возле стартовой точки, связываясь с DPE чуть ниже. После связывания $TF_{II}D$ другие главные факторы транскрипции (для РНК-полимеразы II) формируют основной транскрипционный аппарат на промоторе. Другие элементы промотора, расположенные выше TATA-бокса, связывают активаторы, взаимодействующие с основным аппаратом. Активаторы и основные факторы высвобождаются, когда РНК-полимераза начинает элонгацию.

CTD РНК-полимеразы II фосфорилируется при реакции инициации. И $TF_{II}D$, и белок SRB могут взаимодействовать с CTD. Он также может оказаться местом контакта для белков, модифицирующих РНК-транскрипт, таких как 5′-кэпирующий фермент, факторы сплайсинга и 3′-процессирующий комплекс.

Промоторы могут стимулироваться с помощью энхансеров — последовательностей, способных действовать на значительных расстояниях и в любой ориентации с любой стороны гена. Энхансеры также состоят из ряда консервативных элементов, однако организованных более плотно, чем в промоторах. Некоторые элементы встречаются и в промоторах, и в энхансерах. Вероятно, энхансеры функционируют путем сборки белкового

комплекса, который взаимодействует с белками, связанными на промоторе, что подразумевает «выпетливание» ДНК между ними.

CpG-островки содержат часто встречающиеся CpG-динуклеотиды и обычно окружают промоторы конститутивно экспрессирующихся генов, хотя встречаются также и в промоторах регулируемых генов. Чтобы промотор был способен к инициации транскрипции, такой островок, включая и промотор, должен быть неметилирован. Специальный белок связывается с CpG-динуклеотидами и репрессирует инициацию транскрипции.

Литература

24.2 Эукариотическая ДНК-полимераза состоит из многих субъединиц

Обзор

Doi, R. H. and Wang, L.-F. (1986). Multiple prokaryotic PHK polymerase sigma factors. *Microbiol. Rev.* 50, 227–243.

Young, R. A. (1991). PHK polymerase II. *Annu. Rev. Biochem.* 60, 689–715.

24.4 Промотор РНК-полимеразы I состоит из двух частей

Обзор

Grummt, I. (2003). Life on a planet of its own: regulation of PHK polymerase I transcription in the nucleolus. *Genes Dev.* 17, 1691–1702.

Paule, M. R. and White, R. J. (2000). Survey and summary: transcription by PHK polymerases I and III. *Nucleic Acids Res.* 28, 1283–1298.

Исследование

Bell, S. P., Learned, R. M., Jantzen, H. M., and Tjian, R. (1988). Functional cooperativity between transcription factors UBF1 and SL1 mediates human ribosomal PHK synthesis. *Science* 241, 1192–1197.

24.5 РНК-полимераза III использует и вышележащие, и нижележащие промоторы

Исследование

Bogenhagen, D. F., Sakonju, S., and Brown, D. D. (1980). A control region in the center of the 5S PHK gene directs specific initiation of transcription: II the 3′ border of the region. *Cell* 19, 27–35.

Galli, G., Hofstetter, H., and Birnstiel, M. L. (1981). Two conserved sequence blocks within eukaryotic tPHK genes are major promoter elements. *Nature* 294, 626–631.

Kunkel, G. R. and Pederson, T. (1988). Upstream elements required for efficient transcription of a human U6 PHK gene resemble those of U1 and U2 genes even though a different polymerase is used. *Genes Dev.* 2, 196–204.

Pieler, T., Hamm, J., and Roeder, R. G. (1987). The 5S gene intePHKl control region is composed of three distinct sequence elements, organized as two functional domains with variable spacing. *Cell* 48, 91–100.

Sakonju, S., Bogenhagen, D. F., and Brown, D. D. (1980). A control region in the center of the 5S PHK gene directs specific initiation of transcription: I the 5′ border of the region. *Cell* 19, 13–25.

24.6 $TF_{III}B$ – обязательный фактор для промоторов РНК-полимеразы III

Обзор

Geiduschek, E. P. and Tocchini-Valentini, G. P. (1988). Transcription by PHK polymerase III. *Annu. Rev. Biochem.* 57, 873–914.

Schramm, L. and HePHKndez, N. (2002). Recruitment of PHK polymerase III to its target promoters. *Genes Dev.* 16, 2593–2620.

Исследование

Kassavatis, G. A., Braun, B. R., Nguyen, L. H., and Geiduschek, E. P. (1990). *S. cerevisiae* $TF_{III}B$ is the transcription initiation factor proper of PHK polymerase III, while $TF_{III}A$ and $TF_{III}C$ are assembly factors. *Cell* 60, 235–245.

Kassavetis, G. A., Joazeiro, C. A., Pisano, M., Geiduschek, E. P., Colbert, T., Hahn, S., and Blanco, J. A. (1992). The role of the TATA-binding protein in the assembly and function of the multisubunit yeast PHK polymerase III transcription factor, $TF_{III}B$. *Cell* 71, 1055–1064.

Kassavetis, G. A., Letts, G. A., and Geiduschek, E. P. (1999). A minimal PHK polymerase III transcription system. *EMBO J.* 18, 5042–5051.

24.7 Стартовая точка для РНК-полимеразы II

Обзор

Butler, J. E. and Kadonaga, J. T. (2002). The PHK polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev.* 16, 2583–2592.

Smale, S. T., Jain, A., Kaufmann, J., Emami, K. H., Lo, K., and Garraway, I. P. (1998). The initiator element: a paradigm for core promoter heterogeneity within metazoan protein-coding genes. *Cold Spring Harb Symp Quant Biol.* 63, 21–31.

Smale, S. T. and Kadonaga, J. T. (2003). The PHK polymerase II core promoter. *Annu. Rev. Biochem.* 72, 449–479.

Woychik, N. A. and Hampsey, M. (2002). The PHK polymerase II machinery: structure illuminates function. *Cell* 108, 453–463.

Исследование

Burke, T. W. and Kadonaga, J. T. (1996). *Drosophila* $TF_{II}D$ binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. *Genes Dev.* 10, 711–724.

Singer, V. L., Wobbe, C. R., and Struhl, K. (1990). A wide variety of DNA sequences can functionally replace a yeast TATA element for transcriptional activation. *Genes Dev.* 4, 636–645.

Smale, S. T. and Baltimore, D. (1989). The "initiator" as a transcription control element. *Cell* 57, 103–113.

Weil, P. A., Luse, D. S., Segall, J., and Roeder, R. G. (1979). Selective and accurate initiation of transcription at the Ad2 major late promoter in a soluble system dependent on purified PHK polymerase II and DNA. *Cell* 18, 469–484.

24.8 TBP – универсальный фактор

Обзор

Berk, A. J. (2000). TBP-like factors come into focus. *Cell* 103, 5–8.

HePHKndez, N. (1993). TBP, a universal eukaryotic transcription factor? *Genes Dev.* 7, 1291–1308.

Lee, T. I. and Young, R. A. (1998). Regulation of gene expression by TBP-associated proteins. *Genes Dev.* 12, 1398–1408.

Исследование

Crowley, T. E., Hoey, T., Liu, J. K., Jan, Y. N., Jan, L. Y., and Tjian, R. (1993). A new factor related to TATA-binding protein has highly restricted expression patterns in *Drosophila. Nature* 361, 557–561.

24.9 TBP связывает ДНК необычным способом

Обзор

Burley, S. K. and Roeder, R. G. (1996). Biochemistry and structural biology of TFIID. *Annu. Rev. Biochem.* 65, 769–799.

Lee, T. I. and Young, R. A. (1998). Regulation of gene expression by TBP-associated proteins. *Genes Dev.* 12, 1398–1408.

Orphanides, G., Lagrange, T., and Reinberg, D. (1996). The general transcription factors of PHK polymerase II. *Genes Dev.* 10, 2657–2683.

Исследование

Horikoshi, M. et al. (1988). Transcription factor ATD interacts with a TATA factor to facilitate establishment of a preinitiation complex. *Cell* 54, 1033–1042.

Kim, J. L., Nikolov, D. B., and Burley, S. K. (1993). Cocrystal structure of TBP recognizing the minor groove of a TATA element. *Nature* 365, 520–527.

Kim, Y. et al. (1993). Crystal structure of a yeast TBP/TATA box complex. *Nature* 365, 512–520.

Liu, D. et al. (1998). Solution structure of a TBP-TAFII230 complex: protein mimicry of the minor groove surface of the TATA box unwound by TBP. *Cell* 94, 573–583.

Martinez, E. et al. (1994). TATA-binding protein-associated factors in TFIID function through the initiator to direct basal transcription from a TATA-less class II promoter. *EMBO J.* 13, 3115–3126.

Nikolov, D. B. et al. (1992). Crystal structure of TFIID TATA-box binding protein. *Nature* 360, 40–46.

Ogryzko, V. V. et al. (1998). Histone-like TAFs within the PCAF histone acetylase complex. *Cell* 94, 35–44.

Verrijzer, C. P. et al. (1995). Binding of TAFs to core elements directs promoter selectivity by PHK polymerase II. *Cell* 81, 1115–1125.

Wu, J., Parkhurst, K. M., Powell, R. M., Brenowitz, M., and Parkhurst, L. J. (2001). DNA bends in TATA-binding protein-TATA complexes in solution are DNA sequence-dependent. *J. Biol. Chem.* 276, 14614–14622.

24.10 Основной аппарат собирается на промоторе

Обзор

Nikolov, D. B. and Burley, S. K. (1997). PHK polymerase II transcription initiation: a structural view. *Proc. Natl. Acad. Sci. USA* 94, 15–22.

Zawel, L. and Reinberg, D. (1993). Initiation of transcription by PHK polymerase II: a multi-step process. *Prog. Nucleic Acid Res. Mol. Biol.* 44, 67–108.

Исследование

Buratowski, S., Hahn, S., Guarente, L., and Sharp, P. A. (1989). Five intermediate complexes in transcription initiation by PHK polymerase II. *Cell* 56, 549–561.

Burke, T. W. and Kadonaga, J. T. (1996). *Drosophila* $TF_{II}D$ binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. *Genes Dev.* 10, 711–724.

Bushnell, D. A., Westover, K. D., Davis, R. E., and Kornberg, R. D. (2004). Structural basis of transcription: an PHK polymerase II-TFIIB cocrystal at 4.5 Angstroms. *Science* 303, 983–988.

Littlefield, O., Korkhin, Y., and Sigler, P. B. (1999). The structural basis for the oriented assembly of a TBP/TFB/promoter complex. *Proc. Natl. Acad. Sci. USA* 96, 13668–13673.

Nikolov, D. B. et al. (1995). Crystal structure of a $TF_{II}B$-TBP-TATA-element tePHKry complex. *Nature* 377, 119–128.

24.11 Инициация сопровождается освобождением промотора

Обзор

Calvo, O. and Manley, J. L. (2003). Strange bedfellows: polyadenylation factors at the promoter. *Genes Dev.* 17, 1321–1327.

Hirose, Y. and Manley, J. L. (2000). PHK polymerase II and the integration of nuclear events. *Genes Dev.* 14, 1415–1429.

Price, D. H. (2000). P-TEFb, a cyclin dependent kinase controlling elongation by PHK polymerase II. *Mol. Cell Biol.* 20, 2629–2634.

Proudfoot, N. J., Furger, A., and Dye, M. J. (2002). Integrating mPHK processing with transcription. *Cell* 108, 501–512.

Shilatifard, A., Conaway, R. C., and Conaway, J. W. (2003). The PHK polymerase II elongation complex. *Annu. Rev. Biochem.* 72, 693–715.

Woychik, N. A. and Hampsey, M. (2002). The PHK polymerase II machinery: structure illuminates function. *Cell* 108, 453–463.

Исследование

Douziech, M., Coin, F., Chipoulet, J. M., Arai, Y., Ohkuma, Y., Egly, J. M., and Coulombe, B. (2000). Mechanism of promoter melting by the xeroderma pigmentosum complementation group B helicase of transcription factor IIH revealed by protein-DNA photo-cross-linking. *Mol. Cell Biol.* 20, 8168–8177.

Fong, N. and Bentley, D. L. (2001). Capping, splicing, and 3′ processing are independently stimulated by PHK polymerase II: different functions for different segments of the CTD. *Genes Dev.* 15, 1783–1795.

Goodrich, J. A. and Tjian, R. (1994). Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by PHK polymerase II. *Cell* 77, 145–156.

Holstege, F. C., van der Vliet, P. C., and Timmers, H. T. (1996). Opening of an PHK polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. *EMBO J.* 15, 1666–1677.

Kim, T. K., Ebright, R. H., and Reinberg, D. (2000). Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science* 288, 1418–1422.

Spangler, L., Wang, X., Conaway, J. W., Conaway, R. C., and Dvir, A. (2001). $TF_{II}H$ action in transcription initiation and promoter escape requires distinct regions of downstream promoter DNA. *Proc. Natl. Acad. Sci. USA* 98, 5544–5549.

24.12 Связь между транскрипцией и репарацией

Обзор

Lehmann, A. R. (2001). The xeroderma pigmentosum group D (XPD) gene: one gene, two functions, three diseases. *Genes Dev.* 15, 15–23.

Selby, C. P. and Sancar, A. (1994). Mechanisms of transcription-repair coupling and mutation frequency decline. *Microbiol. Rev.* 58, 317–329.

Исследование

Bregman, D. et al. (1996). UV-induced ubiquitination of PHK polymerase II: a novel modification deficient in Cockayne syndrome cells. *Proc. Natl. Acad. Sci. USA* 93, 11586–11590.

Schaeffer, L. et al. (1993). DNA repair helicase: a component of BTF2 ($TF_{II}H$) basic transcription factor. *Science* 260, 58–63.

Selby, C. P. and Sancar, A. (1993). Molecular mechanism of transcription-repair coupling. *Science* 260, 53–58.

Svejstrup, J. Q. et al. (1995). Different forms of TFIIH for transcription and DNA repair: holo-TFIIH and a nucleotide excision repairosome. *Cell* 80, 21–28.

24.15 Энхансеры несут двунаправленные элементы, способствующие инициации

Обзор

Muller, M. M., Gerster, T., and Schaffner, W. (1988). Enhancer sequences and the regulation of gene transcription. *Eur. J. Biochem.* 176, 485–495.

Исследование

Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of β-globin gene is enhanced by remote SV40 DNA sequences. *Cell* 27, 299–308.

24.16 Энхансеры содержат те же элементы, что и промоторы

Обзор

Maniatis, T., Falvo, J. V., Kim, T. H., Kim, T. K., Lin, C. H., Parekh, B. S., and Wathelet, M. G. (1998). Structure and function of the interferon-beta enhanceosome. *Cold Spring Harbor Symp. Quant. Biol.* 63, 609–620.

Munshi, N., Yie, Y., Merika, M., Senger, K., Lomvardas, S., Agalioti, T., and Thanos, D. (1999). The IFN-beta enhancer: a paradigm for understanding activation and repression of inducible gene expression. *Cold Spring Harbor Symp. Quant. Biol.* 64, 149–159.

24.17 Функция энхансеров состоит в увеличении концентрации активаторов возле промотора

Обзор

Blackwood, E. M. and Kadonaga, J. T. (1998). Going the distance: a current view of enhancer action. *Science* 281, 60–63.

Исследование

Mueller-Storm, H. P., Sogo, J. M., and Schaffner, W. (1989). An enhancer stimulates transcription in *trans* when attached to the promoter via a protein bridge. *Cell* 58, 767–777.

Zenke, M. et al. (1986). Multiple sequence motifs are involved in SV40 enhancer function. *EMBO J.* 5, 387–397.

24.19 CpG-островки являются мишенями для регуляции

Обзор

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev.* 16, 6–21.

Исследование

Antequera, F. and Bird, A. (1993). Number of CpG islands and genes in human and mouse. *Proc. Natl. Acad. Sci. USA* 90, 11995–11999.

Bird, A. et al. (1985). A fraction of the mouse genome that is derived from islands of nonmethylated, Cp-G-rich DNA. *Cell* 40, 91–99.

Boyes, J. and Bird, A. (1991). DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell* 64, 1123–1134.

25

Активация транскрипции

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

25.10 Рецепторы стероидов являются активаторами

- Рецепторы стероидных гормонов относят к группе активаторов, активируемых низкомолекулярными лигандами.
- Связывание лиганда и связывание ДНК осуществляются отдельными доменами.

25.11 Рецепторы стероидов имеют цинковые пальцы

- ДНК-связывающий домен рецептора состоит из цинковых пальцев с аминокислотными остатками Cys вместо стандартных His.
- Рецепторы глюкокортикоидов и эстрогенов имеют по два пальца; первый палец специфичен к последовательности-мишени в ДНК.
- Рецепторы, связывающие ДНК, представляют собой димеры.

25.12 Взаимодействие рецептора со своим лигандом влечет посадку на элемент ответа

- Когда лиганд занимает свой сайт связывания в С-концевом домене рецептора, у рецептора резко повышается специфичное сродство центрального домена к ДНК.

25.13 Рецепторы стероидов узнают свои элементы ответа в соответствии с комбинаторным кодом

- Элемент ответа на стероидный гормон состоит из коротких полусайтов, относящихся друг к другу как инвертированные либо прямые повторы.
- Полусайты бывают всего двух типов.
- Рецептор узнает свой элемент ответа по взаимной ориентации его полусайта и расстоянию, которое их разделяет.
- Последовательность полусайта элемента узнается первым цинковым пальцем рецептора.
- Второй цинковый палец отвечает за димеризацию и, как следствие, за расстояние между субъединицами.
- Расстояние между субъединицами соответствует расстоянию между половинками элементов ответа; от этого зависит способность элемента взаимодействовать с конкретным димером.
- Некоторые рецепторы стероидных гормонов образуют гомодимеры, тогда как другие представители этой группы образуют гетеродимеры.
- Гомодимеры узнают палиндромы, то есть элементы, состоящие из инвертированных половинок; гетеродимеры узнают элементы, состоящие из прямых повторов.

25.14 Гомеодомены узнают в ДНК стандартную последовательность-мишень

- Гомеодомен – это ДНК-связывающий домен из ~60 аминокислот, содержащий три α-спирали.
- Третья, ближняя к С-концу, α-спираль гомеодомена состоит из 17 аминокислот; она укладывается в большую бороздку ДНК.
- N-концевое плечо гомеодомена заходит в малую бороздку ДНК.
- Белки, у которых есть гомеодомен, могут быть не только активаторами, но и репрессорами транскрипции.

25.15 Белки с доменами типа «спираль-петля-спираль» димеризуются в различных сочетаниях

- Домен HLH («спираль-петля-спираль») имеет размер от 40 до 50 аминокислотных остатков; он состоит из двух амфипатических α-спиралей, размером по 15–16 остатков, разделенных петлей – неструктурированным сегментом полипептида.
- Амфипатические спирали ответственны за димеризацию.
- Белки bHLH имеют положительно заряженный мотив, примыкающий к домену HLH; этот мотив отвечает за связывание ДНК.
- В организме животного белки bHLH класса А экспрессируются повсеместно; белки bHLH класса В экспрессируются только в определенных тканях.
- Белки класса В образуют гетеродимеры с белками класса А.
- Белки HLH без положительно заряженного мотива (не bHLH), находясь в гетеродимерах с bHLH, мешают своим партнерам связывать ДНК.
- Белки HLH димеризуются в разных сочетаниях; соотношение этих сочетаний, за счет влияния на паттерн транскрипции, во многом определяет ход индивидуального развития организма.

25.16 «Лейциновые молнии» служат для димеризации

- «Лейциновая молния» объединяет две амфипатических спирали.
- Рядом с «молнией» находятся положительно заряженные участки субъединиц димера; эти участки контактируют с ДНК.
- Димеризация с использованием «лейциновой молнии» рождает домен bZIP; симметричные положительно заряженные участки этого домена связывают инвертированные повторы в ДНК.

25.17 Резюме

25.1 Введение

Фенотипические различия между различными клеточными типами в организме высших эукариот отображают различия в экспрессии генов, кодирующих белки. Эти гены транскрибирует РНК-полимераза II. Теоретически экспрессия таких генов может регулироваться на любом из нескольких возможных этапов. Можно выделить как минимум пять потенциальных «контрольных пунктов», по числу блоков в схеме экспрессии:

Активация структуры хроматина, содержащего ген
↓
Инициация транскрипции
↓
Процессинг транскрипта
↓
Транспорт транскрипта в цитоплазму
↓
Трансляция мРНК в полипептид

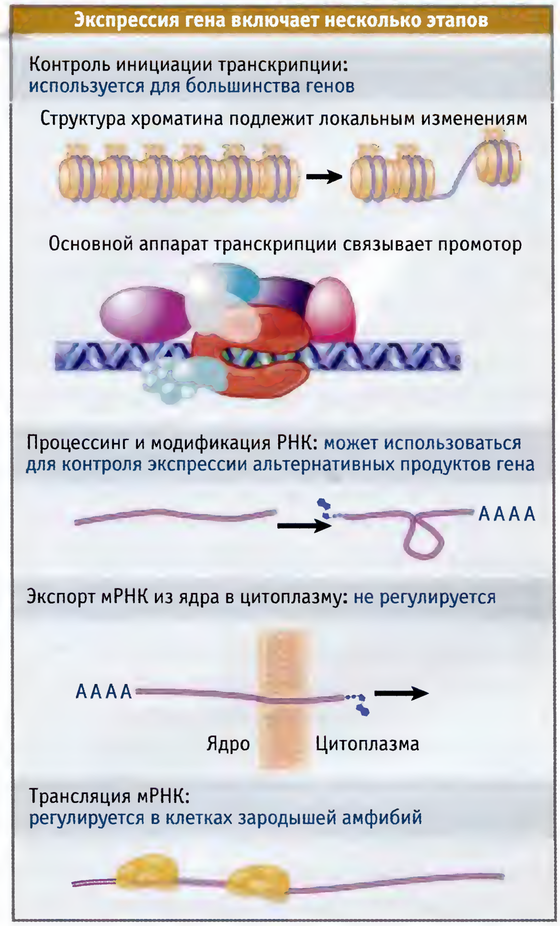


РИС. 25.1 Экспрессия генов контролируется главным образом при инициации транскрипции. Последующие этапы редко влияют на сам факт экспрессии, хотя процессинг мРНК может определять, в какой именно форме экспрессируется ген

В действительности, как видно из РИС. 25.1, экспрессия генов эукариот контролируется, по большей части, при инициации транскрипции. Для большинства генов инициация (в широком смысле — активация) транскрипции — это главный контрольный пункт экспрессии. Чтобы активировать транскрипцию, нужны изменения структуры хроматина в области промотора (см. разд. 30.11 «Активация промотора представляет собой упорядоченную серию событий»). Изменения структуры хроматина сопровождаются посадкой основного аппарата транскрипции (включая саму РНК-полимеразу II) на промотор.

Последующие этапы собственно транскрипции регулируются редко. Для некоторых генов описана преждевременная терминация, которой противодействует киназа P-TEFb; в подавляющем большинстве случаев антитерминация не используется.

Первичный транскрипт гена подлежит модификации: кэпированию 5′-конца и (как правило) полиаденилированию 3′-конца. У транскриптов прерывистых генов вырезаются интроны; после этого зрелая мРНК должна быть экспортирована из ядра в цитоплазму. Регуляция экспрессии генов как отбор транскриптов, чисто теоретически, может коснуться любой стадии процессинга (и даже всех вместе). Однако большинство свидетельств регуляции на уровне процессинга относится к сплайсингу: у некоторых генов есть сразу несколько альтернативных вариантов сплайсинга, и каждому варианту соответствует отдельный белок (конечный продукт экспрессии). О регуляции паттернов сплайсинга читайте в разд. 26.12 «Альтернативный сплайсинг — это реализация различных вариантов сшивания экзонов между собой».

Наконец, трансляция мРНК в цитоплазме тоже может подлежать специфическому контролю. В соматических клетках взрослого организма этот прием используется редко, но в эмбриональном развитии есть моменты, когда регуляция экспрессии генов на уровне трансляции очень актуальна. Такая регуляция может идти путем локализации конкретных мРНК в определенных местах и/или подавления инициации трансляции определенными белковыми факторами.

В многоклеточном эукариотическом организме клеткам свойственно дифференцироваться. В основе любой дифференцировки (происхождения любого типа ткани) лежит регуляция транскрипции генов на этапе активации. Обычно каждый белок-регулятор (фактор транскрипции) обслуживает большое число генов-мишеней. В отношении этого типа регуляции нам надо будет разобраться в двух вопросах: каким образом фактор транскрипции находит (узнает) свою группу генов-мишеней и как зависит активность фактора транскрипции от различных внутренних и внешних сигналов?

25.2 Есть несколько разновидностей факторов транскрипции

Ключевые понятия:

- Основной аппарат транскрипции отвечает за выбор стартовой точки.
- Активаторы задают частоту событий инициации транскрипции.
- Активаторы устанавливают белок-белковые контакты с основными факторами.
- Активаторы могут действовать через коактиваторы.
- Некоторые компоненты аппарата транскрипции действуют путем изменения структуры хроматина.

Инициация транскрипции — это результат множества белок-белковых взаимодействий среди факторов транскрипции, связавших промотор или энхансер; факторы взаимодействуют друг с другом, а также с РНК-полимеразой. Факторы транскрипции делят на несколько классов, перечисленных ниже. Самое общее представление об их свойствах можно составить по РИС. 25.2.

- Основные факторы транскрипции, вместе с РНК-полимеразой, связывают промотор рядом со стартовой точкой и ТАТА-боксом (см. разд. 24.10 «Основной аппарат собирается на промоторе»).
- **Активаторы** — это факторы транскрипции, узнающие специфические сайты (короткие элементы

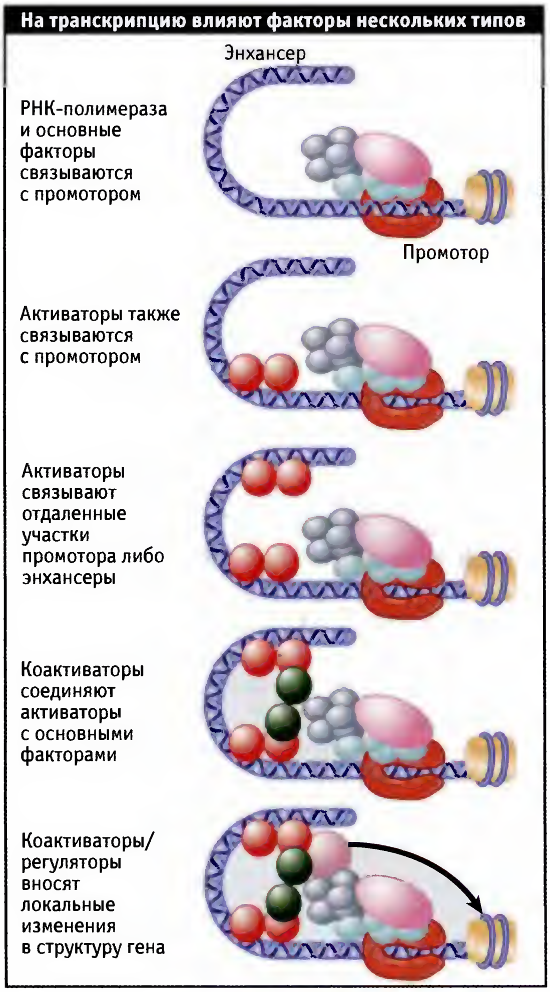


РИС. 25.2 К факторам, участвующим в экспрессии генов, относят РНК-полимеразу и основной аппарат транскрипции, а также активаторы (которые напрямую связывают промоторы либо энхансеры), коактиваторы (связывающиеся с активаторами и в то же время — с основным аппаратом) и регуляторы (воздействующие на структуру хроматина)

промотора или энхансера, объединенные консенсусом). Активаторы связываются со своими сайтами (см. разд. 24.13 «Активаторы связываются с короткими элементами последовательности») и повышают эффективность посадки основного аппарата транскрипции на промотор. Таким образом они повышают частоту событий транскрипции и необходимы для функционирования промотора на адекватном уровне. Некоторые активаторы действуют конститутивно, то есть одинаково во всех клетках. Другим активаторам отводится роль регуляторов; они синтезируются (или активируются) в определенной ткани или в определенное время, контролируя паттерны транскрипции во времени и пространстве. Нуклеотидные последовательности, с которыми связываются такие регуляторы, называют **элементами ответа**.

- Белки следующей группы факторов повышают эффективность транскрипции, не вступая в прямой контакт с ДНК. Это так называемые **коактиваторы**; они служат мостом между активаторами и основным аппаратом транскрипции (см. разд. 25.5 «Активаторы взаимодействуют с основным аппаратом транскрипции»). Действие коактиваторов основано на белок-белковых взаимодействиях.
- Некоторые коактиваторы и целый ряд белков-регуляторов вносят изменения в структуру хроматина (см. разд. 30.7 «Гистонацетилазы ассоциированы с активаторами транскрипции).

Многообразие элементов, из которых можно составить функциональный промотор, а также большое число вариантов расположения этих элементов по отношению к стартовой точке подсказывают, что активаторы как белки взаимодействуют друг с другом множеством различных способов. Расположение элементов в промоторе может быть любым. Промотор можно рассматривать как функциональное сочетание дискретных модулей: об этом говорят результаты экспериментов с перетасовкой участков различных промоторов. Многие гибридные промоторы (например, тимидин-киназы и β-глобина) работают очень хорошо. Значит, элементы ответа в разных промоторах нужны для приведения активаторов в близость с комплексом инициации. Эта близость создает среду для белок-белковых взаимодействий, которые резко повышают эффективность инициации транскрипции.

Организация эукариотических промоторов для РНК-полимеразы II сильно отличается от организации бактериальных промоторов, где все факторы транскрипции должны напрямую взаимодействовать с РНК-полимеразой. В эукариотической системе инициации только основные факторы транскрипции взаимодействуют с ферментом напрямую. Активаторы не взаимодействуют с РНК-полимеразой; они взаимодействуют с основными факторами транскрипции напрямую или опосредованно, через коактиваторы. Такое построение аппарата инициации объясняет большую структурную гибкость промотора (вариабельность расположения элементов друг относительно друга и свободу, с которой они разбросаны в нуклеотидной последовательности генома).

25.3 Связывание ДНК и активация транскрипции осуществляются независимыми доменами фактора

Ключевые понятия:

- **ДНК-связывающая активность и активация транскрипции — это независимые функции независимых доменов активатора.**
- **Задача ДНК-связывающего домена — привлечь активирующий домен к промотору.**

Активаторы и другие регуляторные белки обладают двумя особенностями:

- узнавать специфическую последовательность-мишень в энхансере, промоторе или другой регуляторной области конкретного гена-мишени;
- выполнять собственно активацию посредством связывания с другими компонентами аппарата транскрипции.

Можем ли мы охарактеризовать устройство активатора с точки зрения проявления этих типов активности? Часто у активаторов присутствуют отдельные домены для связывания ДНК и для активации транскрипции. Каждый такой домен — это отдельный функциональный модуль, который функционирует независимо от сочетания с другими доменами. Именно поэтому, принимая во внимание физическую гибкость промотора и подвижную геометрию всего комплекса транскрипции, активирующий домен может контактировать с основным аппаратом транскрипции вне зависимости от точного местоположения и ориентации ДНК-связывающего домена.

Отдельные элементы промотора могут быть удалены от стартовой точки на довольно большое расстояние; они также могут быть ориентированы в любом из двух направлений. Энхансеры могут находиться еще дальше, и их ориентация никогда не имеет значения. Такие свойства регуляторных областей указывают на определенные свойства ДНК и белков-регуляторов. ДНК может образовывать петли или конденсироваться, но этого явно недостаточно. Домены активатора должны быть соединены каким-то гибким способом. Это схематично изображено на РИС. 25.3. Главное, чтобы ДНК-связывающий и активирующий домены активатора не мешали друг другу. Способ их соединения должен позволять активирующему домену взаимодействовать с основным аппаратом безотносительно ориентации и точного местоположения ДНК-связывающего домена.

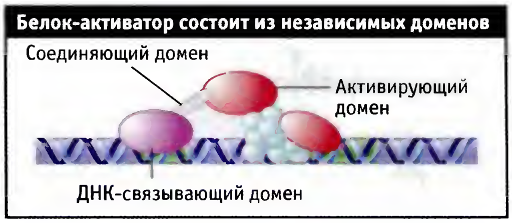


РИС. 25.3 Функции связывания ДНК и собственно активации транскрипции могут выполняться независимыми доменами белка-активатора

Итак, связывание активатора с ДНК необходимо для активации транскрипции. Но зависит ли эффективность активации от наличия в активаторе *конкретного* ДНК-связывающего домена?

Чтобы ответить на этот вопрос, нужно поставить эксперимент, изображенный на РИС. 25.4. У активатора Gal4 есть ДНК-связывающий домен, узнающий элемент *UAS*, в то время как второй, активирующий, домен Gal4 стимулирует инициацию транскрипции с промотора-мишени. В другом белке, бактериальном репрессоре LexA, есть N-концевой ДНК-связывающий домен, узнающий конкретный оператор. Когда LexA садится на этот оператор, он репрессирует промотор, смежный с оператором. В эксперименте по обмену доменами белок Gal4 получает ДНК-связывающий домен LexA вместо своего собственного ДНК-связывающего домена. (Технически это делают так: синтетический ген-матрицу для экспрессии гибридного белка вводят в дрожжевые клетки. Вместе с ним туда помещают ген-репортер, содержащий *UAS* или LexA, и ждут признаков экспрессии гена-репортера.)

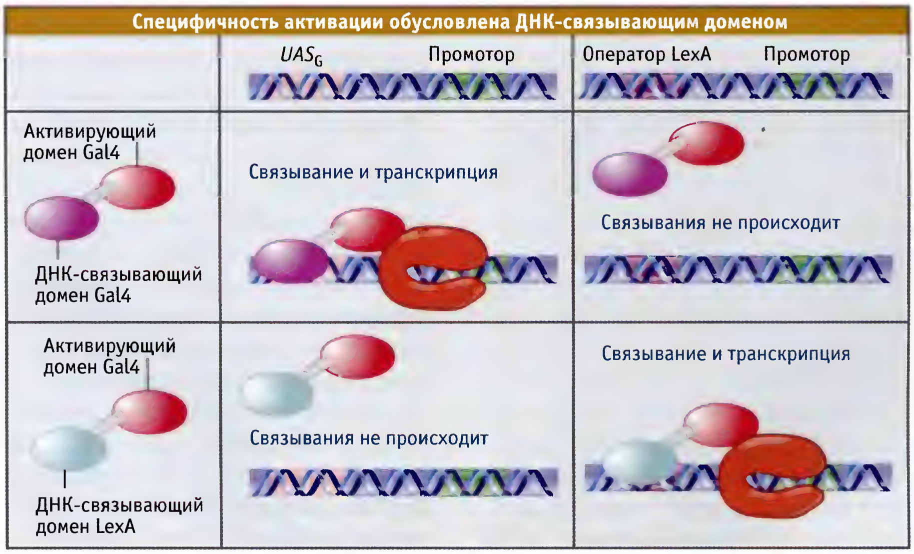


РИС. 25.4 Способность Gal4 активировать транскрипцию не зависит от его специфичности к конкретной ДНК. Если ДНК-связывающий домен Gal4 заменить аналогичным доменом LexA, белок-химера сможет активировать транскрипцию при помощи оператора LexA, помещенного рядом с промотором

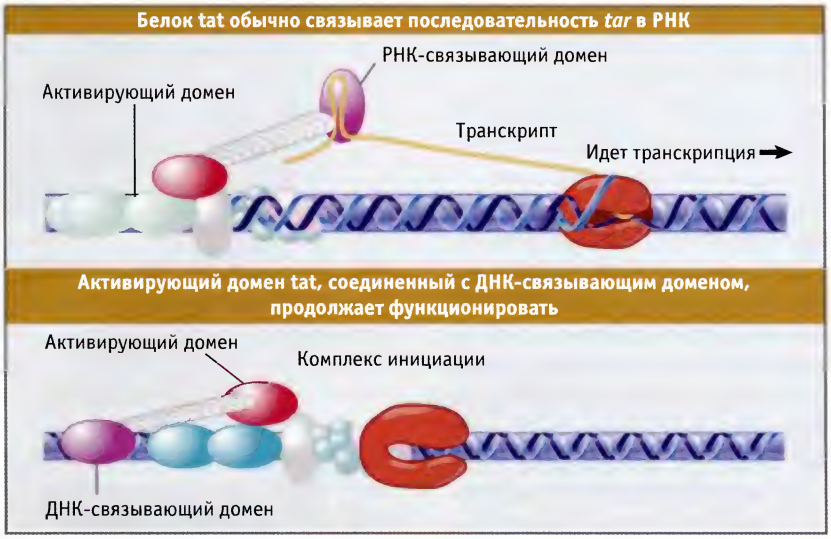


РИС. 25.5 Активирующий домен белка tat вируса иммунодефицита человека (ВИЧ) может стимулировать транскрипцию в том случае, если он удерживается за счет связывания с РНК-продуктом предыдущего раунда транскрипции. На самом деле, способ удержания tat «на рабочем месте» не играет роли: РНК-связывающий домен можно с успехом заменить ДНК-связывающим доменом

Настоящий, природный белок Gal4 может активировать транскрипцию гена-мишени только при наличии *UAS*. Репрессор LexA сам по себе не способен активировать транскрипцию. И наконец, гибридный белок LexA-Gal4 не может активировать транскрипцию из-под *UAS*, зато может активировать транскрипцию из-под оператора LexA.

Этот результат подтверждает концепцию активатора как координатора независимых функциональных модулей. Можно добавить, что ДНК-связывающий домен служит активатору средством доставки к «месту работы». Неважно, в каком месте ДНК-связывающий домен свяжется с ДНК, но там, где он все-таки свяжется, активирующий домен сыграет свою роль. В соответствии с этой моделью средство привлечения активирующего домена к промотору (годится абсолютно любой элемент и подходящий к нему ДНК-связывающий домен) не имеет никакого значения. Легкость создания функциональных гибридных белков путем простого сочетания доменов (функциональных модулей) говорит о том, что при укладке каждый домен обретает активную структуру независимо и без помех со стороны остальной части полипептида.

Представление о том, что связыванием ДНК и собственно активацией транскрипции занимаются разные домены активатора, было развито в исследованиях вируса иммунодефицита человека. Дело в том, что белок tat вируса иммунодефицита человека может стимулировать инициацию транскрипции вообще без связывания с ДНК. Вместо ДНК белок tat связывает продукт транскрипции, РНК, на участке с выраженной вторичной структурой; последовательность этого участка обозначают как *tar* (tat region). На РИС. 25.5 изображена модель стимуляции транскрипции при поддержке взаимодействия tat с *tar*.

Последовательность *tar* находится сразу за стартовой точкой. Поэтому, связывая *tar*, белок tat оказывается в непосредственной близости от комплекса инициации транскрипции. Этого достаточно, чтобы активационный домен tat смог вступить во взаимодействие с одним или несколькими факторами транскрипции в составе комплекса инициации. (Первичный транскрипт синтезируется без участия tat для того, чтобы обеспечить сайт связывания.)

Лучшая демонстрация независимости локализующего и активирующего доменов — это искусственные гибридные белки на основе tat, с заменой *tar*-связывающего мотива ДНК-связывающим доменом. Если вставить в промотор подходящий сайт-мишень, *активационный домен* tat *будет активировать транскрипцию с этого промотора*. Значит, ДНК-связывающий домен действительно является тем «поводком», который удерживает активирующий домен фактора транскрипции вблизи комплекса инициации.

Аналогия с поводком — это частный случай общего представления о том, что для инициации нужна высокая локальная концентрация факторов транскрипции на промоторе. Это достигается, когда активаторы связывают энхансеры в окрестностях промотора, дистальные элементы самого промотора, либо, по крайней мере, 5′-концевую область («верхушку») транскрипта. Все три ситуации рассчитаны на гибкость трехмерной композиции ДНК и белков. Что касается независимости доменов, этот принцип выполняется для всех активаторов транскрипции.

Как мы увидим, функция ДНК-связывающего домена состоит в переносе активационного домена в окрестности стартовой точки. Это объясняет, почему точное местонахождение ДНК-сяхывающих сайтов варьирует в зависимости от промотора.

25.4 Двугибридная система была разработана для поиска новых белок-белковых взаимодействий

Ключевые понятия:

- Двугибридная система позволяет обнаружить белки, взаимодействующие с другими белками. В каждой потенциальной паре один белок искусственно снабжен ДНК-связывающим доменом, другой — активационным доменом.

Принцип независимости доменов активатора лег в основу чрезвычайно полезной разработки. Ее использование привело к открытию огромного числа новых белок-белковых взаимодействий (об этом свидетельствует вал научных публикаций). Метод основан на предположении, что в произвольном активаторе транскрипции есть соединительный домен (вернее, соединительный участок полипептида между двумя функциональными доменами, см. рис. 25.3) и что его можно заменить соединительным белок-белковым взаимодействием.

Предположим, нас интересуют два конкретных белка, X и Y (возможно, не имеющих никакого отношения к транскрипции генов). Могут ли они взаимодействовать между собой? Для ответа на этот вопрос воспользуемся «молекулярным конструктором»: добавим белку X ДНК-связывающий домен (это будет первый гибрид), а белку Y добавим домен, активирующий транскрипцию (это будет второй гибрид). Технически добавление доменов выполняется через слияние сегментов кодирующих последовательностей и экспрессию получившихся гибридных генов.

Если X и Y могут взаимодействовать друг с другом, два гибридных белка тоже смогут взаимодействовать. Этот расчет срабатывает не всегда, но в достаточно большом проценте случаев; поэтому на нем основан метод идентификации белок-белковых взаимодействий, называемый **двугибридной системой**. Представление о том, как работает метод, дает РИС. 25.6.

Белок X, снабженный ДНК-связывающим доменом, связывает сайт (соответствующий этому домену) в промоторе гена-репортера. Самого по себе связывания недостаточно для активации транскрипции, и промотор будет молчать. Исправить ситуацию может гибридный белок Y (если только он действительно взаимодействует с X). Именно взаимодействие X с Y подтянет к промотору активирующий домен.

В качестве репортера походит любой ген с легко детектируемым продуктом. В целом двугибридная система (как набор рутинных автоматизированных процедур) широко используется для быстрого тестирования белок-белковых взаимодействий.

Поразительно: каждый из двух доменов расчлененного активатора транскрипции, попав в чужеродный полипептидный контекст, продолжает очень воспроизводимо и независимо вести себя как полноценный функциональный модуль. Собственно, на модульности белков и основан этот метод. Не менее поразительно то, что способность X и Y к взаимодействию друг с другом не ингибируется присоединением к ним дополнительных функциональных доменов. На практике, конечно, бывает всякое: эффекты нежелательных взаимодействий между доменами и артефакты фолдинга дают себя знать. Любое использование двугибридной системы требует постановки большого числа контрольных экспериментов. И все-таки, в общем и целом, система превосходна.

Еще раз обратим внимание на то, что тестируемые белки, X и Y, не обязаны иметь никакого отношения к транскрипции. Напротив, наши представления о транскрипции (о том, как происходит активация транскрипции) позволили создать новый метод исследования взаимодействия между белками. Если (как мы считаем) ДНК-связывающий и активирующий домены активатора работают независимо, то все, что нужно для активации, — это свести их вместе. Два испытуемых белка сведут их вместе, только взаимодействуя друг с другом в нормальной обстановке клеточного ядра.

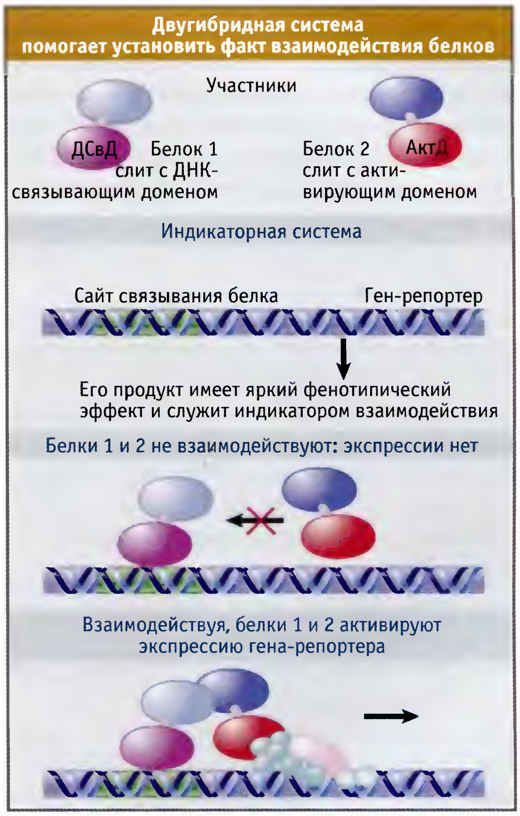


РИС. 25.6 Применение двугибридной системы заключается в создании белковых химер (так называемых гибридов) на основе изучаемых белков и разрозненных доменов фактора транскрипции. Взаимодействие гибридов (говорящее о взаимодействии изучаемых белков) оценивается по их способности совместно активировать транскрипцию

25.5 Активаторы взаимодействуют с основным аппаратом транскрипции

Ключевые понятия:

- Специфичность любого активатора к промотору- или энхансеру-мишени определяется ДНК-связывающим доменом.
- ДНК-связывающий домен отвечает за локализацию домена, активирующего транскрипцию, в непосредственной близости от основного аппарата транскрипции.
- Активатор прямого действия имеет ДНК-связывающий и активирующий домены.
- Активатор, не имеющий активирующего домена, может действовать при посредничестве коактиватора, у которого активирующий домен присутствует.
- В основном аппарате транскрипции есть несколько факторов, которые служат мишенями для активаторов и коактиваторов.

РНК-полимераза, в своем холоферментном комплексе, может быть ассоциирована с различными альтернативными наборами факторов транскрипции. Активатор оказывает прямое действие только в том случае, если он состоит из ДНК-связывающего домена, соединенного с доменом, активирующим транскрипцию (как показано на рис. 25.3). У некоторых активаторов нет своего активирующего домена; вместо этого он связывает еще один белок, коактиватор, у которого такой домен есть. Действие коактиватора изображено на РИС. 25.7. Коактиваторы могут расцениваться в качестве факторов транскрипции, чья специфичность проявляется во взаимодействии с другими факторами транскрипции, но не с ДНК (с которой он вообще не взаимодействует). Конкретному активатору может быть нужен специфический коактиватор.

Для непрямой активации (через коактиватор) действителен все тот же механистический принцип. В то время как активатор прямого действия имеет оба необходимых домена, соединенных ковалентной химической связью, активатор, действующий через коактиватор, по сути, играет роль ДНК-связывающего домена и имеет нековалентные соединения с другими субъединицами (по типу белок-белковых взаимодействий; сравните рис. 25.3 и 25.7). Таким образом, активация транскрипции обусловлена одними и теми же взаимодействиями, независимо от принадлежности функциональных модулей одной субъединице или множеству субъединиц.

Домен, активирующий транскрипцию, делает это за счет установления белок-белковых контактов с основными факторами транскрипции, которые непосредственно управляют сборкой основного аппарата транскрипции. Главная задача активаторов — ускорять сборку основного аппарата. Активатор влияет на основной аппарат благодаря контакту с одним из нескольких основных факторов, чаще всего — с $TF_{II}D$, $TF_{II}B$ или $TF_{II}A$. (Эта тройка основных факторов участвует в ранних этапах сборки основного аппарата, см. рис. 24.14.)



РИС. 25.7 Активатор может связывать коактиватор, контактирующий с основным аппаратом транскрипции

На РИС. 25.8 изображена ситуация, когда контакт уже установлен.

Обычно мишенью активаторов является $TF_{II}D$, в котором есть несколько TAF; контакт может быть установлен с любым из них. На самом деле, факторы TAF главным образом для того и предназначены, чтобы соединять основной аппарат с активаторами. Вот почему сам по себе фактор TBP может поддерживать минимальный уровень транскрипции, а для более интенсивной транскрипции (стимулируемой активаторами) уже нужны факторы TAF из $TF_{II}D$. Разные TAF из состава $TF_{II}D$ могут предоставлять поверхности для взаимодействия с различными активаторами. Некоторые активаторы взаимодействуют только с конкретными TAF, другие — со многими. Полагают, что взаимодействие активаторов с TAF способствует связыванию $TF_{II}D$ с TATA-боксом либо помогает другим активаторам группироваться вокруг комплекса $TF_{II}D$ с TATA-боксом. В любом случае это взаимодействие стабилизирует основной комплекс транскрипции, что ускоряет процесс инициации и повышает мощность промотора (частоту событий инициации и интенсивность транскрипции).

У дрожжей активирующие домены активаторов Gal4 и Gcn4 несут множество отрицательных зарядов; за это их называют «кислыми активаторами». Еще один исключительно эффективный кислый активатор — белок VP16

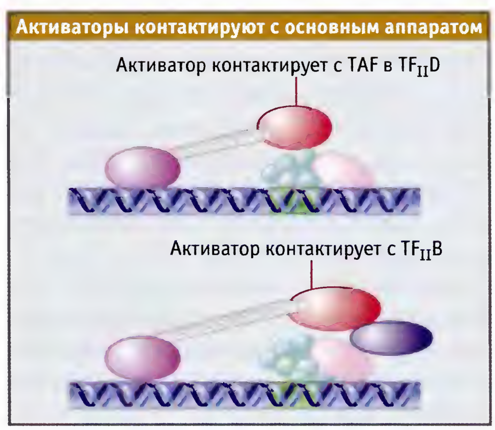


РИС. 25.8 Активаторы могут действовать на разных стадиях инициации, контактируя с субъединицами TAF в $TF_{II}D$ либо с $TF_{II}B$

вируса герпеса. У VP16 нет своего ДНК-связывающего домена, так что формально его следует отнести к коактиваторам. Многие эксперименты по описанию функций кислого активатора проводились на VP16 с «пришитым» ДНК-связывающим доменом.

Кислые активаторы усиливают способность другого основного фактора, $TF_{II}B$, к вступлению в основной комплекс инициации. Эксперименты, проведенные *in vitro*, показали, что связывание $TF_{II}B$ с комплексом инициации на промоторе аденовируса стимулируется в присутствии кислого активатора (Gal4 или VP16), причем VP16 может связаться прямо с $TF_{II}B$. Можно заключить, что лимитирующим этапом инициации с этого промотора является вступление $TF_{II}B$ в основной комплекс и что кислый активатор его стимулирует.

Промоторы для РНК-полимеразы II не очень чувствительны к перестановкам и даже заменам отдельных элементов. Это говорит о низкой специфичности активации транскрипции с этих промоторов. Такое впечатление, что любой активатор, чей активирующий домен окажется в пределах досягаемости основного комплекса инициации, сумеет содействовать его сборке. Яркое подтверждение такой универсальности было получено при испытании искусственных промоторов, содержащих новые произвольные комбинации элементов. Например, если поместить элемент UAS_G, взятый у дрожжей, рядом с промотором гена мыши, этот ген начнет активироваться кислым активатором Gal4 в культивируемых клетках млекопитающих. Получается, что средства активации промотора, которыми располагает активатор Gal4, подходят и дрожжам, и высшим эукариотам (то есть эти средства высококонсервативны). По всей видимости, белок Gal4 распознает в основном аппарате транскрипции нечто общее для дрожжей и млекопитающих.

Каким же образом активатор стимулирует транскрипцию? Напрашивается два типа моделей:

- Согласно рекрутинговой модели, единственный эффект активатора заключается в повышении сродства РНК-полимеразы к промотору (recruitment model).
- В альтернативной модели активатор индуцирует какое-то изменение в комплексе транскрипции (например, в конформации фермента), повышающее его эффективность.

При тестировании этих моделей оказалось, что у дрожжей при активации повышается сродство РНК-полимеразы. Однако после того, как концентрация РНК-полимеразы была сильно увеличена, активатор больше не мог повышать эффективность транскрипции. По всей видимости, чистый вклад активатора все-таки состоит в повышении эффективной концентрации РНК-полимеразы на промоторе.

Если собрать вместе все компоненты, требуемые для эффективной транскрипции (основные факторы, РНК-полимеразу, активаторы и коактиваторы), получится очень большой аппарат из > 40 белков. Возможно ли, чтобы такой аппарат проходил поэтапную сборку на промоторе? Оказывается, некоторые активаторы, коактиваторы и основные факторы могут проходить начальные этапы сборки прямо на промоторе; после этого к ним присоединяется огромный комплекс: РНК-полимераза, уже укомплектованная другими активаторами и коактиваторами. Все это показано на РИС. 25.9.

На сегодня известно несколько белковых комплексов, в которых РНК-полимераза ассоциирована с различными факторами транскрипции. Самый внушительный из них, «холоферментный комплекс» у дрожжей, состоит из РНК-полимеразы, ассоциированной с двадцатисубъединичным комплексом, так называемым **медиатором**. Считается, что «холоферментный комплекс» способен инициировать транскрипцию без помощи дополнительных компонентов. В состав медиатора входят продукты экспрессии нескольких генов, инактивация которых блокирует транскрипцию, например продукты локусов *SRB* (название связано с тем, что многие гены в этих локусах были идентифицированы по супрессии мутаций в гене РНК-полимеразы В; suppressor of РНК polymerase B). Название «медиатор» было выбрано из-за способности комплекса быть посредником в действии активаторов. Медиатор необходим дрожжам для транскрипции большинства генов. У высших эукариот есть гомологичные комплексы, также необходимые для транскрипции большинства генов. Когда медиатор взаимодействует с С-концевым доменом РНК-полимеразы II, он претерпевает изменение конформации. Медиатор может передавать РНК-полимеразе эффекты активаторов и репрессоров. Когда РНК-полимераза начинает элонгациию транскрипта, медиатор, скорее всего, уходит. Некоторые факторы транскрипции прямо влияют на транскрипцию, взаимодействуя с РНК-полимеразой или с основным аппаратом. Другие факторы влияют косвенно, через контроль структуры хроматина (см. разд. 30.3 «Ремоделирование хроматина — это активный процесс»).

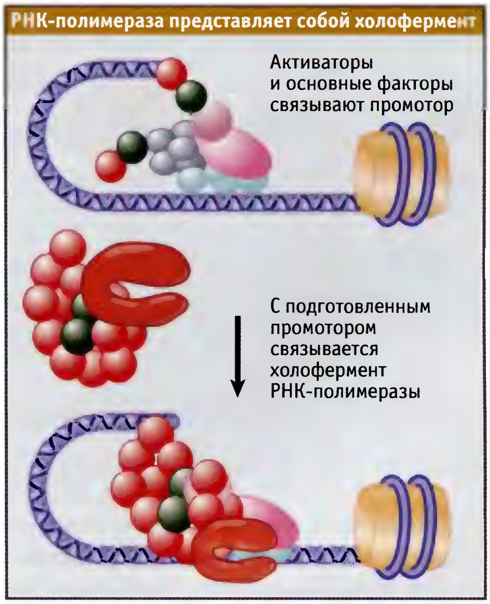


РИС. 25.9 РНК-полимераза представляет собой холофермент, содержащий множество активаторов

25.6 Некоторые связывающие промотор белки являются репрессорами

Ключевые понятия:

- В большинстве случаев репрессия транскрипции достигается за счет изменения структуры хроматина на обширной области генома; лишь у некоторых промоторов есть свои, специфичные к ним, белки репрессоры.

Обычно репрессия транскрипции генов эукариот выполняется на уровне структуры хроматина. Хотя белки-регуляторы, блокирующие транскрипцию по типу репрессоров в бактериальных оперонах, в эукариотах встречаются довольно редко, несколько примеров все же известно. Репрессоры такого типа играют активную роль в подавлении функции основного аппарата транскрипции. Одним из примеров служит репрессор NC2/Dr1/DRAP1. Он представляет собой гетеродимер, связывающий TBP и не дающий ему взаимодействовать с другими компонентами основного аппарата транскрипции. О важности этого репрессора можно судить по нежизнеспособности дрожжей, у которых полностью инактивирован любой из генов, кодирующих компоненты репрессора.

В другом, более частном, примере мишенью регуляции служит CAAT-бокс. У морского ежа в промоторе гена гистона H2B, нарабатываемого только в ходе сперматогенеза, есть две копии CAAT-бокса (см. рис. 24.22). В семенниках, где идет сперматогенез, CAAT-связывающие

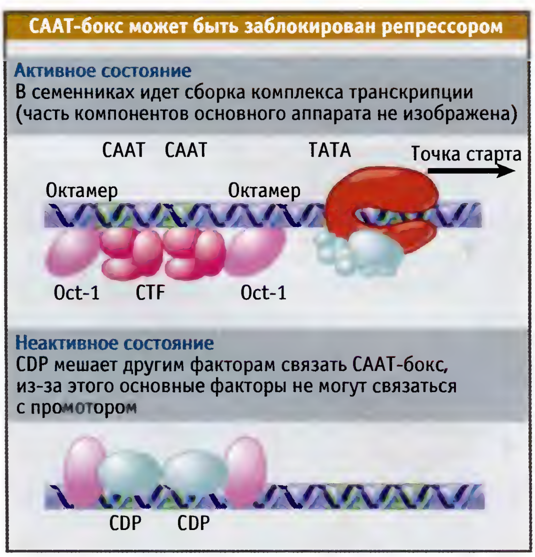


Рис. 25.10 Сборка комплекса транскрипции в семенниках морского ежа требует опознания ряда элементов промотора H2B. Связывание фактора, блокирующего один из этих элементов, CAAT-бокс, мешает связыванию фактора, узнающего именно этот элемент, и, таким образом, препятствует образованию активного комплекса

факторы (CAAT-binding factor, CTF) связывают CAAT-бокс и активируют транскрипцию. В остальных тканях присутствует другой белок — репрессор CDP (CAAT-displacement protein); связывая CAAT-боксы, тем самым *закрывает их от активатора*, предотвращая их узнавание.

На рис. 25.10 изображены последствия такой регуляции. В семенниках различные элементы промотора гена H2B (CAAT-боксы, TATA-бокс, октамеры) все вместе обеспечивают связывание факторов транскрипции. В остальных тканях эмбриона удаление CAAT-связывающего фактора с промотра препятствует образованию транскрипционного комплекса. В данном случае действие CDP аналогично действию бактериального репрессора, который не позволяет РНК-полимеразе инициировать транскрипцию с промотора, на котором он сидит. Случай с CDP демонстрирует, что, если какой-нибудь белок связывает конкретный, пусть даже хорошо изученный, элемент промотора, о функции этого белка можно только гадать: он может быть активатором, репрессором или влиять на транскрипцию с этого промотора.

25.7 Элементы ответа узнаются активаторами

Ключевые понятия:

- Элементы ответа могут входить в состав промоторов и энхансеров.
- Каждый такой элемент узнается специфичным к нему активатором.
- В промоторе или энхансере может быть много элементов ответа; они могут участвовать в активации транскрипции как поодиночке, так и в определенных комбинациях.

Общая закономерность, которую можно сформулировать, исходя из описаний различных групп генов с общими системами контроля, выглядит так: *в промоторах или энхансерах генов каждой группы есть идентичные сегменты ДНК, опознаваемые активатором*. Сегмент промотора или энхансера, заставляющий ген отвечать на присутствие фактора, называют **элементом ответа** [(response element; например, элемент ответа на тепловой шок (heat shock response element, HSE); элемент ответа на глюкокортикоиды (glucocorticoid response element, GRE); элемент ответа на сыворотку (serum response element, SRE)]. Копии «элементов ответа», относящиеся к разным генам, очень похожи между собой (то есть не обязательно полностью идентичны), и в них обязательно есть короткие консенсусы. Сегмент ДНК, связываемый фактором, с обеих сторон немного выступает за края консенсуса. В промоторах элементы ответа находятся не на фиксированном расстоянии от стартовой точки, но обычно в пределах 200 пн от нее. Для того чтобы получить регуляторный ответ в форме всплеска экспрессии, обычно достаточно одной копии элемента, но существуют промоторы, у которых присутствует несколько копий.

Элементы ответа могут находиться как в промоторах, так и в энхансерах. Некоторые типы элементов характерны именно для промоторов, другие — в основном для энхансеров. Например, HSE чаще всего попадаются в промоторах, а GRE — в энхансерах. Предположительно, все элементы ответа функционируют по одному и тому же принципу: связывание активатора с элементом ответа необходимо для того, чтобы РНК-полимераза могла начать транскрипцию гена. Отличие активаторов, работающих с элементами ответа, от конститутивных активаторов состоит в том, что они работают (присутствуют в активной форме) только в определенных условиях (возможно, нестандартных для клетки). Момент, когда ген начнет экспрессироваться, зависит от наступления этих условий.

Примером нестандартных условий, в которых много генов контролируется одним и тем же фактором, служит повышение температуры окружающей среды, приводящее клетку в состояние теплового шока. Очень многие клетки (даже бактериальные) умеют реагировать на повышение температуры, регулируя экспрессию множества своих генов по особой схеме, известной как ответ на тепловой шок (heat shock response). В рамках ответа транскрипция некоторых генов выключается; транскрипция целого ряда других генов, наоборот, активируется (это так называемые гены теплового шока). Кроме изменений в транскрипции генов повышение температуры вызывает изменения в трансляции мРНК. На примере регуляции экспрессии «генов теплового шока можно продемонстрировать различия в способах контроля экспрессии у прокариот и эукариот. У бактерий синтезируется новый σ-фактор, настраивающий холофермент РНК-полимеразы на узнавание альтернативной последовательности (~10 пн), присутствующей в промоторах генов теплового шока (см. разд. 11.16 «Инициация на данном промоторе завсит от типа сигма-фактора»). В ядре эукариотической клетки гены теплового шока тоже содержат схожие консенсусные последовательности (HSE), но, во-первых, эти элементы находятся на разных расстояниях от стартовых точек; во-вторых, с ними работает независимый активатор HSTF — транскрипционный фактор теплового шока (не являющийся ни компонентом основного аппарата транскрипции, ни, тем более, субъединицей РНК-полимеразы). Таким образом, активация фактора HSTF служит средством прицельной избирательной активации транскрипции ~20 генов, в промоторах которых есть соответствующая последовательность-мишень.

У *Drosophila melanogaster* все гены теплового шока снабжены множеством копий HSE, потому что ее HSTF обладает свойством кооперативно связывать смежные копии элемента. Поразительно: в ходе эволюции HSE оставались настолько консервативными, что ген теплового шока из *D. melanogaster* может быть активирован со своего собственного промотора в клетках морского ежа или мыши. Что касается белков HSTF, они тоже крайне консервативны; у плодовой мушки и у дрожжей они очень похожи и оставляют одинаковые молекулярные «отпечатки» на ДНК, несущей элементы ответа на тепловой шок. У дрожжей транскрипционный фактор подлежит фосфорилированию: именно так, при наступлении теплового шока, происходит его активация.

Примером того, как один и тот же ген может оказаться на пересечении множества регуляторных контуров, служит экспрессия гена металлотионеина. Белок металлотионеин, МТ, защищает клетку от избыточных концентраций тяжелых металлов: он связывает ион металла и выводит его из клетки. Ген МТ, который в стандартных условиях экспрессируется на некотором уровне, может быть индуцирован к повышенной экспрессии; первичными индукторами служат ионы тяжелых металлов (например, кадмия) либо глюкокортикоиды. Регуляторная область гена МТ сочетает в себе элементы ответа нескольких типов.



РИС. 25.11 Регуляторная область гена металлотионеина человека содержит регуляторные элементы как в промоторе, так и в энхансере. В промоторе есть элементы, через которые идет индукция экспрессии металлотионеина ионами металлов; в энхансере есть элемент ответа на глюкокортикоиды. Под картой промотора с названиями элементов на рисунке обозначены белки, связывающие эти элементы

Организация промотора гена MT схематично изображена на РИС. 25.11. Сразу бросается в глаза, что элементы, через которые может идти активация транскрипции, в этом промоторе расположены очень плотно. TATA- и GC-боксы занимают свои обычные положения неподалеку от стартовой точки. Еще два элемента, BLE (от *англ.* basal level elements — элементы базового уровня) формально подходят под определение энхансеров. Хотя они расположены не очень далеко от стартовой точки, их можно переместить гораздо дальше без ущерба для функции. Кроме того, BLE содержат последовательности, характерные для энхансеров; их связывают те же белки, что и энхансер SV40.

Элемент ответа на эфиры форболовых кислот (из-за которых в тканях образуются опухоли) обозначается как TRE (TPA response element). Его консенсус есть в нескольких энхансерах, например в одном из BLE гена MT, а также в 72 пн повторах вируса SV40. Элемент TRE содержит сайт связывания фактора AP1; в стандартных условиях это связывание работает как часть механизма конститутивной экспрессий, одним из активаторов которой служит AP1. В других условиях у взаимодействия AP1 с TRE другая функция: оно служит одним из рычагов запуска изменений в паттерне транскрипции, направленных на нейтрализацию последствий атаки форболовых эфиров.

Индуктивный ответ гена MT на присутствие тяжелых металлов происходит благодаря наличию в его промоторе множества элементов MRE (metal response element; см. рис. 25.11). Наличие в промоторе любого гена хотя бы одного MRE придает гену способность реагировать повышенной экспрессией на присутствие тяжелых металлов. Чем больше таких элементов, тем интенсивнее индукция; с MRE взаимодействует фактор транскрипции MTF1.

Еще один элемент в промоторе MT обеспечивает ответ на стероидные гормоны: это GRE, лежащий в 250 пн от стартовой точки. GRE ведет себя как энхансер. Делеция участка, содержащего GRE, не влияет ни на базальный уровень экспрессии, ни на индукцию экспрессии MT ионами тяжелых металлов. Зато без GRE не происходит никакого ответа на стероиды.

Регуляция экспрессии металлотионеина хорошо иллюстрирует тот общий принцип, согласно которому *любой из нескольких различных элементов в энхансере или промоторе может независимо активировать ген.* Отсутствие элемента, через который проходит один путь активации, никак не влияет на исправность остальных путей активации. Многообразие элементов, их независимость друг от друга, практически неограниченная свобода их расположения говорят о том, что фактор, связавший любой из этих элементов, может независимо, как подобает активатору, влиять на основной аппарат транскрипции. Это влияние, основанное на белок-белковых взаимодействиях, стабилизирует аппарат транскрипции или способствует его сборке, таким образом во много раз повышая частоту событий инициации.

25.8 Различные типы ДНК-связывающих доменов

Ключевые понятия:

- **Активаторы принято классифицировать по типу ДНК-связывающего домена.**
- **Внутри одного класса активаторов возможны вариации аминокислотной последовательности функционально значимого мотива, определяющие специфичность разных белков к конкретным сайтам-мишеням.**

Активаторам свойственно иметь модульную структуру. За связывание с ДНК и за активацию транскрипции в них отвечают два отдельных белковых домена. Чаще всего активаторы классифицируют по типу ДНК-связывающего домена. На самом деле, за связывание с ДНК отвечает лишь относительно короткий мотив этого домена.

- Мотив **«цинковый палец»** (zinc finger) был впервые идентифицирован у фактора $TF_{III}A$, который требуется РНК-полимеразе III для транскрипции генов 5S рРНК. С тех пор «цинковые пальцы» были обнаружены у многих факторов транскрипции (безусловных и предполагаемых). Особой разновидностью этого мотива обладают рецепторы стероидных гормонов.
- **Рецепторы стероидных гормонов** (steroid receptors) объединены в одну группу не за общий мотив, а за общий механизм их активации. Каждый из рецепторов стероидных гормонов — это, на самом деле, классический активатор; просто для связывания со своим элементом в ДНК этому активатору необходимо сначала связать конкретный стероид. Рецепторы глюкокортикоидов, вместе с рецептором тиреоидных гормонов, рецептором ретиноевой кислоты и некоторыми другими внутриклеточными рецепторами, объединяют в активаторов, активируемых лигандами (ligand-activated activators). Всем им свойствен один и тот же *modus operandi: белковый фактор не проявляет своей активности до тех пор, пока не свяжет низкомолекулярный лиганд.*
- Мотив **«спираль-поворот-спираль»** (helix-turn-helix) был впервые идентифицирован у бактериофагов, в ДНК-связывающем домене лизогенного репрессора. В классическом варианте, одна α-спираль лежит в большой бороздке ДНК, другая α-спираль, располагаясь под углом к первой, ложится поперек ДНК. Похожий мотив есть в **гомеодоменах**. Гомеодомены были впервые обнаружены у *Drosophila* в нескольких белках, признанных важнейшими регуляторами эмбрионального развития. Похожие мотивы присутствуют во многих факторах транскрипции млекопитающих.
- Амфипатический мотив **«спираль-петля-спираль»** (helix-loop-helix, HLH) был идентифицирован в других белках-регуляторах эмбрионального развития и во многих других эукариотических ДНК-связывающих белках. Амфипатическая спираль с двух сторон

образует поверхности с противоположными свойствами. На одной стороне собраны гидрофобные аминокислотные остатки, на другой стороне находятся положительно и отрицательно заряженные гидрофильные остатки; размер соединительной петли составляет 12–28 аминокислотных остатков. Этот мотив позволяет белкам образовывать димеры; контакт с ДНК осуществляет другой участок полипептида, соседний с мотивом: он содержит много положительно заряженных аминокислотных остатков.

- **«Лейциновые молнии»** образованы особыми участками полипептидов, где в каждом седьмом положении стоит лейцин. Каждый такой участок — это половинка молнии; он взаимодействует со своей половинкой из другого полипептида. Получается димер, «застегнутый на молнию». Рядом с каждой половинкой молнии группируются положительно заряженные аминокислотные остатки, обеспечивающие димеру контакт с ДНК.

Активность индуцибельного активатора может регулироваться несколькими различными путями, как это схематично изображено на РИС. 25.12.

- Фактор имеет пространственную специфичность, поскольку синтезируется не во всех, а только в определенных клетках или тканях. Самый яркий пример — регуляторы развития типа белков с гомеодоменами у *Drosophila*.
- Активность фактора может зависеть от внесения в него определенной химической модификации. Например, транскрипционный фактор теплового шока HSTF переводится в активную форму фосфорилированием. Фактор AP1, представляющий собой гетеродимер Jun и Fos, переходит в активную форму в результате фосфорилирования субъединицы Jun.
- Фактор активируется (либо инактивируется) лигандом. Яркий пример — рецепторы стероидов. Связывание лиганда может влиять на локализацию фактора (вызывая его перемещение из цитоплазмы в ядро) либо придавать фактору способность к связыванию ДНК.
- Фактор не всегда доступен; например, известнейший фактор NF-κB, который, в частности, активирует гены иммуноглобулинов κ в В-лимфоцитах представлен во многих клетках. В большинстве клеток фактор

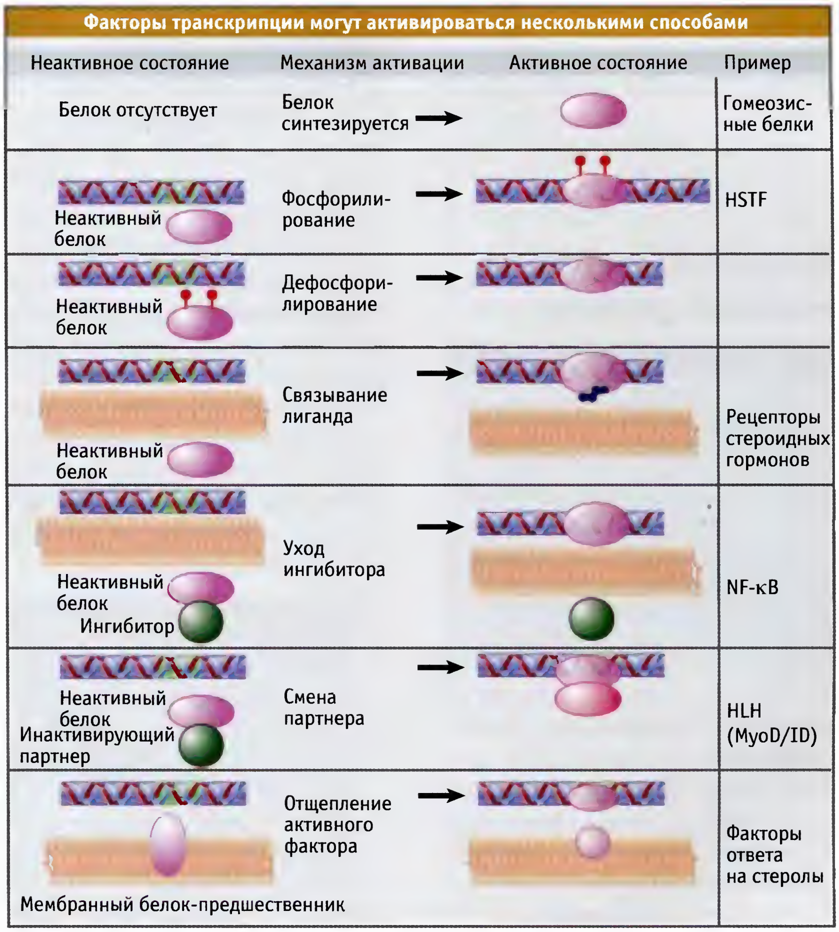


РИС. 25.12 Активность регуляторного фактора транскрипции может контролироваться на уровне синтеза белка, ковалентной модификации белка, связывания фактора с лигандом, а также за счет действия ингибиторов, которые секвестрируют фактор (не пускают белок в целевой компартмент) либо нарушают его способность связывать ДНК

удерживается в цитоплазме за счет взаимодействия с белком-ингибитором I-κB. В B-лимфоцитах NF-κB освобождается от I-κB; фактор перемещается из цитоплазмы в ядро и там активирует транскрипцию.

- У фактора-димера могут быть альтернативные комбинации субъединиц, особенно характерные для белков с доменами HLH. Например, предположим, комбинация субъединиц A и B неактивна; синтез C, альтернативного партнера A, приведет к вытеснению B из комплекса AB с образованием активного комплекса AC. Это элементарная схема; в принципе такие схемы могут переплетаться между собой: множество альтернативных партнеров образуют множество альтернативных комбинаций друг с другом. Соотношение комбинаций определяет характер регуляции.
- Активный фактор может быть получен при протеолитическом расщеплении белка-предшественника. Например, существует активатор, который вначале представляет собой белок, прикрепленный к ядерной оболочке или эндоплазматическому ретикулуму со стороны цитоплазмы. Отсутствие в мембранах стеролов (например, холестерола) приводит к отщеплению цитоплазматического домена, который направляется в ядро и там работает в качестве активной формы активатора.

Далее мы детально рассмотрим события связывания ДНК и реакцию активации в исполнении активаторов из разных классов.

25.9 «Цинковые пальцы» — это ДНК-связывающие домены

Ключевые понятия:

- «Цинковый палец» — это мотив из ~23 аминокислотных остатков; он имеет вид петли, основание которой скреплено атомом цинка, контактирующим с остатками гистидинов и цистеинов.
- У белка с «цинковыми пальцами» этих «пальцев», как правило, несколько.
- У каждого «пальца» C-концевая часть имеет вид α-спирали, которая ложится в большую бороздку ДНК.
- Некоторые белки с «цинковыми пальцами» связывают РНК (вместо ДНК либо наряду с ДНК).

«Цинковые пальцы» — это особая разновидность доменной организации белка, при которой небольшие консервативные сегменты полипептида связывают по одному иону цинка; при этом каждый сегмент работает как единичный (и в определенной степени независимый) функциональный мотив. Классический пример белков с «цинковыми пальцами» — стероидные рецепторы.

Белок с «цинковыми пальцами», как правило, имеет целую серию таких «пальцев»; типовая серия «цинковых пальцев» изображена на РИС. 25.13. Аминокислотный консенсус одного «пальца» записывается так:

$$\text{Cys-X}_{2-4}\text{-Cys-X}_3\text{-Phe-X}_5\text{-Leu-X}_2\text{-His-X}_3\text{-His}$$

Само название «цинковый палец» представляет собой тривиальное описание петли, выступающей из сайта связывания цинка на поверхности белковой глобулы. Мотиву также подходит альтернативное название «Cys_2/His_2-палец», потому что при ближайшем рассмотрении ион цинка удерживается внутри тетраэдра с вершинами из двух консервативных цистеинов и двух консервативных гистидинов. В последовательность «пальца» входит 23 аминокислотных остатка; промежуток между пальцами обычно равняется семи-восьми остаткам.

«Цинковые пальцы» есть у многих ДНК-связывающих белков. В первичной структуре полипептида домен из «пальцев» можно идентифицировать как серию тандемных повторов (обычно такая серия одна, но бывает и больше). У белка ADR1 *D. melanogaster* этот домен очень маленький: в нем всего два «пальца». У $TF_{III}A$, напротив, много «пальцев»: целых девять, и на них уходит почти весь полипептид. ДНК-связывающий домен известного активатора Sp1 состоит из трех «цинковых пальцев».

Анализ кристаллической структуры ДНК, состоящей в комплексе с «трехпалым» белком, послужил основой для модели, изображенной на РИС. 25.14. В соответствии с этой моделью C-концевая часть каждого «пальца» имеет вид α-спирали, контактирующей с ДНК. N-концевые части «пальцев» образуют β-слой. Для простоты, этот β-слой, а также положения атомов цинка, никак не обозначены на схеме в нижней части рисунка. Три α-спирали (от трех «пальцев») помещаются в один виток большой бороздки ДНК. Каждая α-спираль узнает в ДНК определенную нуклеотидную последовательность и образует с ней контакты в двух точках, отмеченных на схеме стрелками. Считается, что вариабельные аминокислотные остатки в C-концевых частях «пальцев» обеспечивают более узкую специфичность к конкретным последовательностям-мишеням.

Поскольку «цинковые пальцы» были обнаружены у многих истинных активаторов (стимулирующих активность РНК-полимераз II и III), любой новый белок с пальцами претендует, прежде всего, на роль фактора транскрипции. Этот подход полностью оправдал



РИС. 25.13 Фактор транскрипции SP1 имеет серию из трех «цинковых пальцев». Аминокислотная последовательность каждого «пальца» характеризуется особым расположением цистеинов и гистидинов, образующих сайт связывания иона цинка

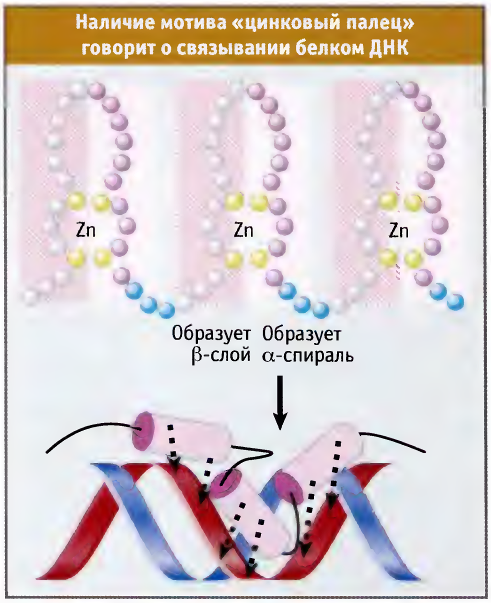


РИС. 25.14 «Цинковые пальцы» могут принимать форму α-спиралей, которые входят в большую бороздку ДНК, смыкаясь с β-слоями по другую сторону двойной спирали

себя во многих исследованиях; в частности, он помог отождествить несколько локусов *D. melanogaster* с функциями важнейших регуляторов эмбрионального развития.

На самом деле, в интерпретации наличия у белка «цинковых пальцев» необходимо соблюдать осторожность (особенно в неоднозначных ситуациях с сомнительным консенсусом или в случаях наличия одинарного мотива). Как выяснилось, «пальцы» могут связывать РНК вместо ДНК либо вообще не иметь никакого отношения к связыванию нуклеиновых кислот. Например, типовой белок с «цинковыми пальцами», $TF_{III}A$, связывает не только ген 5S, но и продукт этого гена, 5S рРНК. Фактор инициации трансляции eIF2β тоже имеет «цинковый палец»; мутации, нарушающие аминокислотную последовательность этого пальца, влияют на эффективность и специфичность узнавания инициирующих кодонов. Наконец, капсидные белки ретровирусов тоже содержат мотив, похожий на «палец»; предполагают, что этот «палец» каким-то образом взаимодействует с вирусной РНК.

25.10 Рецепторы стероидов являются активаторами

Ключевые понятия:

- Рецепторы стероидных гормонов относят к группе активаторов, активируемых низкомолекулярными лигандами.
- Связывание лиганда и связывание ДНК осуществляются отдельными доменами.

В организме животного стероидные гормоны (стероиды; небольшие органические молекулы) синтезируются в ответ на различные нейроэндокринные стимулы; эти гормоны сильнейшим образом влияют на процессы роста, дифференцировки и состояние гомеостаза у высших животных. Классификация стероидов (и других соединений с похожими молекулярными активностями) приведена на РИС. 25.15.

Важным источником стероидов являются надпочечники. Они секретируют более 30 разновидностей стероидов; в основном это глюкокортикоиды и минералокортикоиды. Половые стероидные гормоны (мужские, андрогены, и женские, эстрогены) синтезируются в специализированных репродуктивных органах и тканях. Витамин D, по своей структуре родственный стероидам, необходим для формирования костной ткани.

Существуют гормоны, которые, имея другую, чем стероиды, химическую природу и другой, чем у стероидов, физиологический смысл, по механизму своего действия очень похожи на стероиды: например, тиреоидные гормоны (йодированные производные тирозина), которые контролируют общую интенсивность метаболизма. У ряда таксонов стероидные и тиреоидные гормоны (например, экдистероиды у насекомых и тиреоидные гормоны у амфибий) играют важные роли в таком переломном этапе жизненного цикла, как метаморфоз.

Действие ретиноевой кислоты, или витамина А, может иметь множество разнообразных физиологических смыслов. Например, в эмбриональном развитии птиц ретиноевая кислота служит морфогеном, градиент которого размечает передне-заднюю ось будущей конечности. Метаболит ретиноевой кислоты, 9-*цис*-ретиноевую кислоту, можно обнаружить в тканях, где идет интенсивное накопление и переработка витамина А.

Многообразие эффектов гормона (или другой сигнальной молекулы) отображает многообразие путей регуляции, в которых могут быть задействованы эти гормоны. Различные химические соединения, которым посвящен этот раздел, характеризуются общим способом воздействия: *в качестве низкомолекулярных лигандов они специфично связывают свои рецепторы, побуждая их активировать транскрипцию генов-мишеней.*

В привычном понимании, рецептор — это белок, закрепленный в клеточной мембране и способный взаимодействовать с лигандом на поверхности клетки. Белки, которые принято называть рецепторами стероидных (тиреоидных) гормонов, имеют другие свойства. В данном случае сам термин рецепторы был подобрано не совсем удачно, поскольку рецептор стероидного или тиреоидного гормона является рецептором ровно настолько, насколько репрессор оперона *lac* у *E. coli* является рецептором β-галактозида.

Рецепторы различных групп стероидных гормонов, а также рецепторы тиреоидных гормонов и рецепторы ретиноевой кислоты принадлежат к особому классу регуляторов экспрессии — так называемых активаторов, активирующихся лигандом (ligand-responsive activators). У этих «рецепторов» связывание лиганда и связывание

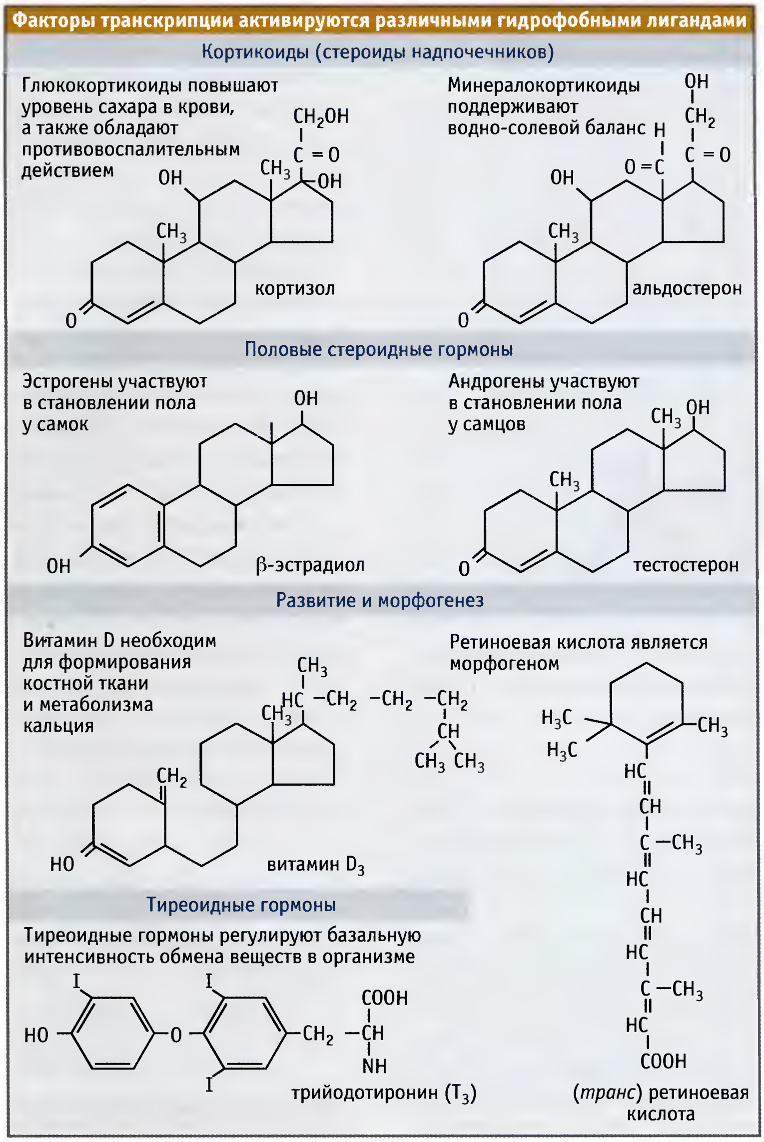


РИС. 25.15 **Несколько разновидностей гидрофобных низкомолекулярных соединений работают как активаторы факторов транскрипции**

ДНК осуществляются независимыми доменами, расположенными друг относительно друга стандартным образом. Большое структурное сходство различных активаторов этого класса показано на РИС. 25.16.

ДНК-связывающий домен образован центральной частью полипептида. Эти регионы очень похожи у различных рецепторов (совпадет от 42% до 94% аминокислот). Надо сказать, что у активаторов этого класса акт связывания ДНК неотделим от собственно активации (поскольку мутации в ДНК-связывающем центральном домене могут нарушать обе эти функции).

N-концевые области полипептидов, напротив, очень вариабельны. Они помогают центральному домену активировать транскрипцию. С-концевые домены рецепторов занимаются связыванием гормонов. В рецепторах стероидов аминокислотные последовательности этих доменов совпадают на 30–57%; различия отражают специфичность рецепторов к конкретным гормонам. С другими рецепторами этого класса сходство прослеживается слабо, в силу специфичности этих рецепторов к совсем другим лигандам: витамину D, тиреоидным гормонам, ретиноевой кислоте и т.д. В этом же С-концевом домене также есть мотивы, обеспечивающие димеризацию рецепторов, мотив, задействованный в активации транскрипции.

Есть лиганды, для которых существует множество близкородственных рецепторов; например у ретиноевой кислоты есть рецепторы RARα, RARβ и RARγ, а у 9-*цис*-ретиноевой кислоты — рецепторы RXRα, RXRβ и RXRγ.



РИС. 25.16 Рецепторы многих стероидных и тиреоидных гормонов имеют похожую организацию, с особой N-концевой областью, консервативной ДНК-связывающей областью, и С-концевой областью, предназначенной для связывания гормона. Сравнение проведено относительно рецептора глюкокортикоидов

25.11 Рецепторы стероидов имеют «цинковые пальцы»

Ключевые понятия:

- ДНК-связывающий домен «рецептора» состоит из «цинковых пальцев» с аминокислотными остатками Cys вместо стандартных His.
- Рецепторы глюкокортикоидов и эстрогенов имеют по два «пальца»; первый «палец» специфичен к последовательности-мишени в ДНК.
- «Рецепторы», связывающие ДНК, представляют собой димеры.

Рецепторы стероидных гормонов (а также некоторые другие белки) имеют мотивы «цинковые пальцы», отличающиеся от стандартных «пальцев» Cys_2/His_2. Их структура основывается на альтернативном консенсусе, тоже приспособленном к удержанию атома цинка:

$$Cys\text{-}X_2\text{-}Cys\text{-}X_{13}\text{-}Cys\text{-}X_2\text{-}Cys$$

Такие Cys_2/Cys_2 «пальцы» есть у многих «рецепторов», причем нередко в единственном числе (как мы помним, у белков со стандартными Cys_2/His_2 «пальцами» этих «пальцев» как минимум два). Все известные последовательности-мишени (сайты связывания) «рецепторов» в ДНК представляют собой короткие прерывистые палиндромы (то есть пару коротких инвертированных повторов, разделенных небольшим интервалом).

Рецепторы глюкокортикоидов и эстрогенов имеют по два «пальца»; в каждом «пальце», по центру тетраэдра из цистеинов, содержится атом цинка. Совместный фолдинг α-спиралей, образуемых «пальцами», дает большой глобулярный домен. Гидрофобный центр этой глобулы составляют ароматические поверхности α-спиралей, совместно с β-слоем, соединяющим две спирали. Одна из сторон α-спирали первого (считая от N-конца полипептида) «пальца» имеет контакты в большой бороздке ДНК.

Чтобы связать ДНК, рецептор димеризуется. Каждой субъединице (половинке) димера достается один из смежных витков двойной спирали ДНК. Манера рецепторов димеризоваться находится в полном соответствии с симметричной структурой элементов ответа на стероиды (и им подобные сигнальные молекулы; см. разд. 25.13 «Рецепторы стероидов узнают элементы ответа в соответствии с комбинаторным кодом»).

Каждый «палец» отвечает за одну из принципиально важных функциональных характеристик рецептора; функционально значимые аминокислотные остатки отмечены на РИС. 25.17. Несколько остатков на правой стороне первого «пальца» диктуют нуклеотидную последовательность мишени (элемента ответа в ДНК). Несколько остатков на левой стороне второго «пальца» задают интервал между симметричными половинками мишени (каждая половинка мишени узнается одной из двух субъединиц димера рецептора; см. разд. 25.13 «Рецепторы стероидов узнают элементы ответа в соответствии с комбинаторным кодом»).

Факт связывания ДНК первым «пальцем» был убедительно доказан в экспериментах по обмену специфичностями. У рецептора эстрогенов посредством делеции удаляли «палец» и заменяли его гомологичным «пальцем», взятым у рецептора глюкокортикоидов. Получившийся белок-химера узнавал в ДНК элемент GRE (стандартную мишень рецептора глюкокортикоидов) и больше не узнавал ERE (стандартную мишень рецептора эстрогенов). Так было доказано, что специфичность узнавания «рецептором» мишени в ДНК зависит от аминокислотной



РИС. 25.17 Первый «цинковый палец» рецептора стероидов определяет, какая нуклеотидная последовательность будет связана этим рецептором (значимые аминокислоты отмечены фиолетовым цветом). Второй «палец» задает интервал между сайтами связывания (значимые аминокислоты отмечены синим цветом)

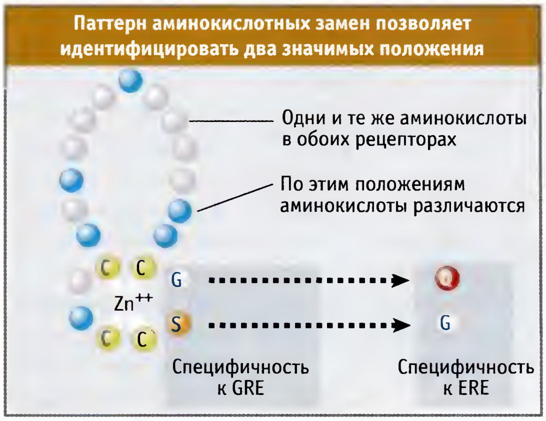


РИС. 25.18 Узнавание последовательностей GRE и ERE как отличных друг от друга обусловливается двумя аминокислотами в основании первого «цинкового пальца» рецептора

последовательности первого «цинкового пальца» этого рецептора.

У двух рассмотренных рецепторов, глюкокортикоидного и эстрогенового, различия этой последовательности сосредоточены в основании первого «пальца». На самом деле, замена аминокислотных остатков всего в двух (особо важных) положениях, отмеченных на РИС. 25.18, уже позволяет рецептору глюкокортикоидов связывать ERE вместо GRE.

25.12 Взаимодействие рецептора со своим лигандом влечет посадку на элемент ответа

Ключевые понятия:

- Когда лиганд занимает свой сайт связывания в С-концевом домене рецептора, у рецептора резко повышается специфичное сродство центрального домена к ДНК.

На РИС. 25.19 изображено хорошо изученное взаимодействие глюкокортикоидов со своим рецептором, а также показан эффект этого взаимодействия. Стероидный гормон легко попадает в клетку: он проникает сквозь наружную мембрану за счет простой диффузии. В клетке гормон находит свой рецептор и связывается с ним. Локализация свободных рецепторов точно не установлена. Возможно, их концентрация уравновешена между ядром и цитоплазмой. Когда гормон связывает свой рецептор, тот переходит в активированную форму с повышенным сродством к ДНК; поэтому комплекс гормон-рецептор всегда обнаруживается в ядре.

Активированный рецептор глюкокортикоидов узнает в ДНК конкретный нуклеотидный консенсус (GRE). Большинство GRE находится в энхансерах; как известно, энхансер может быть отдален от промотора в любую сторону на расстояние до нескольких тпн. Тем не менее в ответ на связывание GRE комплексом гормон-рецептор энхансер эффективно активирует промотор: с него начинает идти транскрипция. Связывание рецептором своего элемента в энхансере — это типичный для стероидов путь регуляции экспрессии огромного множества генов.

Активность различных рецепторов зависит от состояния С-концевого домена. Если у рецептора глюкокортикоидов удалить С-концевой домен, получится *конститутивно активный* рецептор; он будет активировать транскрипцию через GRE *вне зависимости от присутствия лиганда*. По всей видимости, в отсутствие глюкокортикоидов, С-концевой домен их рецептора мешает рецептору узнавать GRE; этот домен как бы служит рецептору внутренним отрицательным регулятором. С приходом лиганда ингибирование ослабевает; теперь рецептор эффективно связывает GRE, и транскрипция активируется. Можно пойти еще дальше: предположить, что ингибирование на самом деле не такое уж внутреннее, но выполняется за счет взаимодействия С-концевого домена с каким-нибудь особенным белком, который уходит при появлении лиганда.

У рецептора эстрогенов все по-другому: если у него удалить домен, ответственный за связывание лиганда, рецептор по-прежнему будет связывать ERE, но уже безо всякого эффекта для транскрипции. Следовательно, С-концевой домен рецептора эстрогенов работает не как пассивный внутренний ингибитор (инактивируемый лигандом), а как внутренний активатор (комплектуемый лигандом).

Осталось добавить, что большая часть результатов, обсуждаемых в этом разделе, была получена с использованием синтетического аналога глюкокортикоидов, гормона дексаметазона.

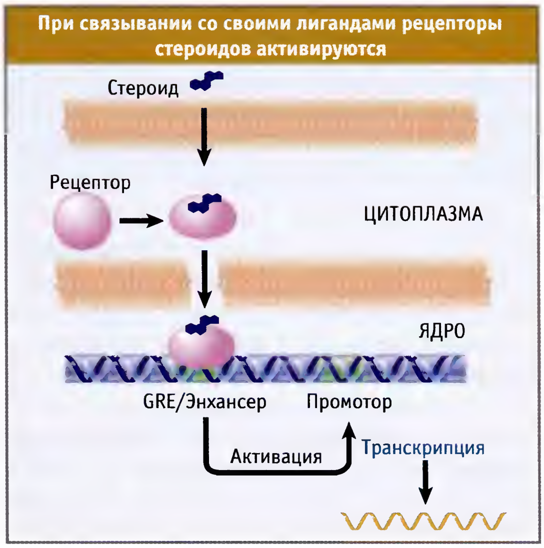


РИС. 25.19 Глюкокортикоиды стимулируют транскрипцию генамишени, заставляя свои рецепторы связывать энхансер, необходимый для функционирования промотора

25.13 Рецепторы стероидов узнают элементы ответа в соответствии с комбинаторным кодом

Ключевые понятия:

- Элемент ответа на стероидный гормон состоит из коротких полусайтов, относящихся друг к другу как инвертированные либо прямые повторы.
- Полусайты бывают всего двух типов.
- Рецептор узнает свой элемент ответа по взаимной ориентации его полусайтов и расстоянию, которое их разделяет.
- Последовательность полусайта элемента узнается первым «цинковым пальцем» рецептора.
- Второй «цинковый палец» отвечает за димеризацию и, как следствие, за расстояние между субъединицами димера с ДНК.
- Расстояние между субъединицами соответствует расстоянию между половинками элементов ответа; от этого зависит способность элемента взаимодействовать с конкретным димером.
- Некоторые рецепторы стероидных гормонов образуют гомодимеры, тогда как другие представители этой группы образуют гетеродимеры.
- Гомодимеры узнают палиндромы, то есть элементы, состоящие из инвертированных половинок; гетеродимеры узнают элементы, состоящие из прямых повторов.

Каждый рецептор узнает в ДНК элемент ответа, состоящий из двух коротких повторов, или полусайтов (half sites). Это предполагает, что активной формой фактора, взаимодействующего с таким элементом, должен быть димер (у которого каждая субъединица взаимодействует со своей половинкой элемента). С подобной ситуацией мы уже сталкивались при разборе взаимодействия лизогенного репрессора фага лямбда со своим оператором в разд. 14.11 «Мотив «спираль-поворот-спираль» необходим репрессору для связывания ДНК».

Полусайты (половинки элемента ответа) на стероидный гормон могут иметь одну и ту же ориентацию (прямой повтор) либо быть инвертированы друг относительно друга. В промоторе или энхансере эти полусайты могут быть разнесены на 0–4 пн (причем последовательность этого промежутка значения не имеет). Для всего множества рецепторов существует всего два типа полусайтов. Специфичность рецепторов к конкретным элементам определяется расположением полусайтов: их взаимной ориентацией в сочетании с расстоянием, которое их разделяет. Впрочем, правила узнавания не являются непреложными, но до определенной степени зависят от контекста; к тому же известны случаи пермиссивного узнавания одной и той же пары инвертированных повторов двумя-тремя разными рецепторами.

Весь класс так называемых рецепторов (регуляторов транскрипции, запускаемых от взаимодействия с лигандами) можно поделить на две группы:

- К первой группе относят рецепторы глюкокортикоидов (GR), минералокортикоидов (MR), андрогенов (AR) и прогестерона (PR). Все эти рецепторы образуют гомодимеры (то есть димеры, состоящие из одинаковых субъединиц). Они узнают в ДНК элементы, состоящие из половинок с консенсусом TGTTCT. На РИС. 25.20 показано, что эти половинки ориентированы как палиндром (то есть они инвертированы друг относительно друга). Узнавание такого элемента конкретным рецептором (димером) полностью зависит от расстояния (промежутка) между половинками. Рецептор эстрогенов (ER) тоже входит в эту группу, но стоит в ней несколько особняком: он тоже узнает сайт из инвертированных половинок, но с другим консенсусом (TGACCT вместо TGTTCT).
- Рецептор 9-*цис*-ретиноевой кислоты (RXR) образует не только гомодимеры, но и гетеродимеры с ~15 другими рецепторами, среди которых есть рецепторы тиреоидных гормонов (T3R), витамина D (VDR) и ретиноевой кислоты (RAR). На РИС. 25.21 показано, что такие димеры узнают половинки элементов с консенсусом TGACCT. Половинки представляют собой прямые повторы; при узнавании промежуток между ними всегда учитывается. Для активации рецептора-гетеродимера достаточно связывания лиганда с одной из субъединиц. Одни гетеродимеры активируются лигандом партнера RXR, другие могут быть активированы любым из двух лигандов. Многие партнеры RXR образуют еще и гомодимеры, узнающие в ДНК элементы-палиндромы.

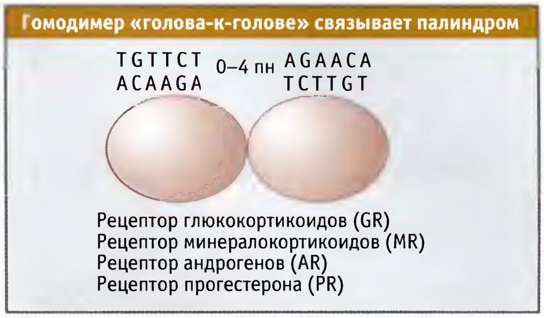


РИС. 25.20 Элементы, образуемые инвертированными половинками сайта (TGTTCT), узнаются несколькими разными рецепторами в зависимости от интервала между этими половинками



РИС. 25.21 Элементы с прямыми повторами TGACCT узнаются гетеродимерами с участием RXR

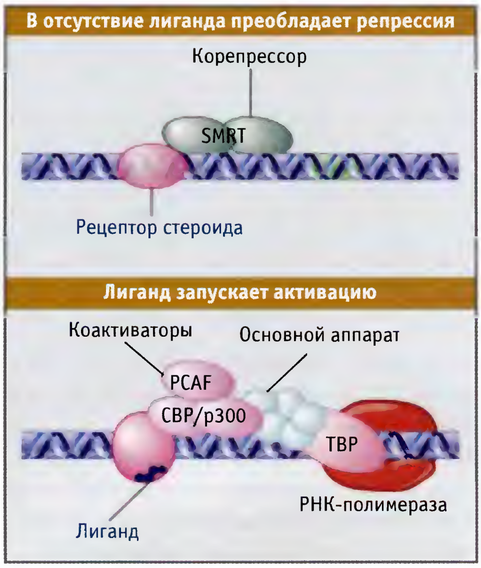


РИС. 25.22 В отсутствие лиганда рецепторы стероидов, TR и RAR, связывают корепрессор SMRT. Промотор не функционирует. Когда из-за связывания лиганда со своим рецептором SMRT приходится уйти, рецептор связывается с комплексом-коактиватором, что, в конечном счете, приводит к активации основного аппарата транскрипции

Теперь мы вплотную подошли к пониманию моторики регуляторных манипуляций в исполнении целой армии стероидов и родственных низкомолекулярных агентов. Вспомним рис. 25.17; там было показано, что узнавание консенсуса в полусайте элемента зависит от *первого* «цинкового пальца» в С-концевом домене, а узнавание расстояния между половинками — от аминокислотной последовательности *второго* «пальца» в С-концевом домене. Этот «палец» задает такие особенности строения димера, от которых зависит расстояние между субъединицами, сидящими в смежных витках большой бороздки ДНК. Таким образом, второй «палец» определяет соответствие рецептора определенному размеру промежутка между половинами элемента ответа. Положение аминокислот, ответственных за гетеродимеризацию, специфично для каждой пары мономеров.

Каким же образом рецепторы стероидных гормонов активируют транскрипцию? Как выяснилось, у них нет прямой связи с основным аппаратом транскрипции. При этом их обслуживает очень мощный коактиватор — целый комплекс со своим набором активностей. Один из компонентов этого комплекса, CBP/p300 модифицирует структуру хроматина за счет ацетилирования гистонов (и это только одна из функций CBP/p300; см. рис. 30.14).

Еще раз напомним, что все белки («рецепторы») рассматриваемого класса работают как активаторы транскрипции, функция которых зависит (активизируется) от взаимодействия с лигандами. Оказывается, некоторые представители данного класса в определенных условиях способны еще и к подавлению транскрипции. Выяснилось, что рецепторы TR и RAR, находясь в гетеродимерах с RXR и в отсутствие лиганда, в некоторых локусах работают как репрессоры. Выполнение такого рода функции обеспечивается способностью гетеродимеров взаимодействовать с белком-корепрессором. Принцип действия корепрессора противоположен принципу действия коактиваторов: корепрессор всячески мешает работе основного аппарата транскрипции (в частности, за счет деацетилирования гистонов, см. рис. 30.16). Какой реальный вес (в физиологических условиях ответа на гормональную стимуляцию) может иметь такая репрессирующая активность рецепторов» против их лиганд-зависимой активации, пока неясно.

Если свободный рецептор репрессирует транскрипцию, связывание рецептора с лигандом нейтрализует репрессию. Более того, рецептор, укомплектованный лигандом, немедленно становится активатором. В отсутствие лиганда рецептор взаимодействует с комплексом-корепрессором (он напрямую контактирует с одним из компонентов корепрессора, SMRT). Связывание лиганда вызывает у рецептора конформационное изменение, приводящее к разрыву контакта с SMRT. Это изменение позволяет коактиватору занять место корепрессора; как показано на РИС. 25.22, это полностью меняет характер регуляции.

25.14 Гомеодомены узнают в ДНК стандартную последовательность-мишень

Ключевые понятия:

- Гомеодомен — это ДНК-связывающий домен из ~60 аминокислот, содержащий три α-спирали.
- Третья, ближняя к С-концу, α-спираль гомеодомена состоит из 17 аминокислот; она укладывается в большую бороздку ДНК.
- N-концевое плечо гомеодомена заходит в малую бороздку ДНК.
- Белки, у которых есть гомеодомен, могут быть не только активаторами, но и репрессорами транскрипции.

Гомеобокс — это особая нуклеотидная последовательность, присутствующая в кодирующей области некоторых генов. Гомеобокс транслируется как участок полипептида размером ~60 аминокислотных остатков; при фолдинге этот участок неизменно складывается в домен с определенной структурой, — так называемый гомеодомен. Белки с гомеодоменом есть у подавляющего большинства (если не у всех) эукариотических организмов.

Название «гомеобокс» было дано в связи с историей первоначальной идентификации подобных последовательностей в гомеозисных локусах *Drosophila*. Уже тогда было известно, что работа этих локусов задает первич-

ную спецификацию отделов тела в развитии зародыша мухи (как мы теперь знаем, не только мухи, но и слона). С той поры похожие мотивы были обнаружены в генах самых разных эукариотических организмов (в частности, во многих генах-регуляторах эмбрионального развития). Аминокислотные мотивы, напоминающие гомеодомен, можно вычленить во многих факторах транскрипции из разных таксонов животных.

В продуктах гомеозисных генов *Drosophila* (полипептидах) гомеодомен чаще всего (но не всегда) расположен недалеко от С-конца. На РИС. 25.23 приведены примеры генов, в которых есть гомеобокс. Такие гены могут быть совсем не похожи между собой, если не считать близкого родства гомеобоксов. На самом деле, гомеобоксы тоже бывают разные (консервативность — понятие условное). Например, самая главная группа гомеобокс-содержащих генов *Drosophila* очень упорядочена: у любых двух генов этой группы гомеобоксы совпадают как минимум на 80%. Но это, скорее, исключение: гомеобоксы из разных генов (тем более у разных организмов) могут довольно сильно отличаться. Ноконец, во многих факторах транскрипции классический гомеодомен может сопровождаться разными аминокислотными мотивами. Типичным примером служат октамер-связывающие белки Oct, у которых с областью, напоминающей классический гомеодомен, соседствует так называемый Pou домен из ~75 аминокислот.

Гомеодомен отвечает за связывание с ДНК; эксперименты по созданию белковых химер из разных доменов показали, что специфичность связывания заложена в самом гомеодомене. Но, как и во многих других случаях (например, в случае с репрессорами бактериофагов), не существует простого соответствия между аминокислотной последовательностью гомеодомена и нуклеотидной последовательностью его мишени в ДНК. С-концевая часть гомеодомена имеет значительное сходство с мотивом «спираль-поворот-спираль»; такой мотив есть в некоторых прокариотических репрессорах. Мы помним (это обсуждалось в разд. 14.11 «Мотив «спираль-поворот-спираль» необходим репрессору для связывания ДНК»),



РИС. 25.23 Гомеодомен может быть единственным ДНК-связывающим доменом белка-регулятора транскрипции или присутствовать в комбинации с другими аналогичными мотивами. Гомеодомен выражен как дискретная часть белка размером примерно 60 аминокислот

Гомеодомен – это модуль из 60 аминокислот

	1		N-концевое плечо							10				Спираль 1						20		
En	Glu	Lys	Arg	Pro	Arg	Thr	Ala	Phe	Ser	Ser	Glu	Gln	Leu	Ala	Arg	Leu	Lys	Arg	Glu	Phe	Asn	Glu
Antp	Arg	Lys	Arg	Gly	Arg	Gln	Thr	Tyr	Thr	Arg	Tyr	Gln	Thr	Leu	Glu	Leu	Glu	Lys	Glu	Phe	His	Phe
Oct-2	Arg	Arg	Lys	Lys	Arg	Thr	Ser	Ile	Glu	Thr	Asn	Val	Arg	Phe	Ala	Leu	Glu	Lys	Ser	Phe	Leu	Ala

							30				Спираль 2							40
En	Asn	Arg	Tyr	Leu	Thr	Glu	Arg	Arg	Arg	Glu	Glu	Leu	Ser	Ser	Glu	Leu	Gly	Leu
Antp	Asn	Arg	Tyr	Leu	Thr	Arg	Arg	Arg	Arg	Ile	Glu	Ile	Ala	His	Ala	Leu	Cys	Leu
Oct-2	Asn	Glu	Lys	Pro	Thr	Ser	Glu	Glu	Ile	Leu	Leu	Ile	Ala	Glu	Gln	Leu	His	Met

	41									50				Спираль 3						60
En	Asn	Glu	Ala	Gln	Ile	Lys	Ile	Trp	Phe	Gln	Asn	Lys	Arg	Ala	Lys	Ile	Lys	Lys	Ser	Asn
Antp	Thr	Glu	Arg	Gln	Ile	Lys	Ile	Trp	Phe	Gln	Asn	Arg	Arg	Met	Lys	Trp	Lys	Lys	Glu	Asn
Oct-2	Glu	Lys	Glu	Val	Ile	Arg	Val	Trp	Phe	Cys	Asn	Arg	Arg	Gln	Lys	Glu	Lys	Arg	Ile	Asn

РИС. 25.24 Гомеобокс *Antennapedia* типичен для самой известной группы гомеобокс-содержащих генов *Drosophila*; ген *engrailed* (*en*) принадлежит к другому классу гомеозисных генов; фактор Oct-2 млекопитающих относится к классу факторов транскрипции, родственных продуктам гомеозисных генов. Гомеодомен начинается N-концевым плечом, за которым следуют три спиралевидных мотива, образованных аминокислотными остатками 10-22, 28-38 и 42-58. На рисунке аминокислоты гомеодомена пронумерованы от 1 до 60, а консервативные аминокислоты (совпадающие у всех трех белков) напечатаны синим шрифтом

что лизогенный репрессор фага лямбда имеет одну α-спираль, специфичную к нуклеотидной последовательности сайта связывания репрессора в ДНК оператора. Эта α-спираль (спираль-3, или «спираль узнавания») входит в большую бороздку ДНК и обеспечивает контакты; в это время другая α-спираль репрессора, α-спираль-2, ложится на ДНК снаружи под некоторым углом. Похожим образом, в виде трех α-спиралей, может быть организован и гомеодомен. Во всяком случае, в первичной структуре гомеодоменов хорошо просматриваются три потенциальных α-спирали; сравнения проведены на РИС. 25.24. Третья спираль приходится на самую консервативную часть гомеодомена; правда, по сравнению с репрессором фага лямбда, у которого третья спираль образована всего-навсего девятью аминокислотными остатками, гомеодомен имеет почти вдвое более длинную третью спираль.

Структура гомеодомена белка engrailed *D. melanogaster* схематично изображена на РИС. 25.25. Большинство контактов между белком и нуклеиновой кислотой обеспечивает спираль-3, входящая в большую бороздку ДНК. Вообще многие контакты, удерживающие и ориентирующие этот мотив в большой бороздке ДНК, связывают только сахаро-фосфатный остов двойной спирали ДНК; такие контакты не могут быть специфичны к нуклеотидной последовательности ДНК. Эти контакты приходятся в основном на одну из сторон двойной спирали и фланкируют те несколько азотистых оснований ДНК, на которые приходятся *специфичные* контакты. Кроме третьей α-спирали взаимодействовать с ДНК гомеодомену помогает его N-концевое плечо; оно заходит в малую бороздку. Таким образом, за взаимодействие с ДНК отвечают в первую очередь С-концевой и N-концевой аминокислотные мотивы гомеодомена.

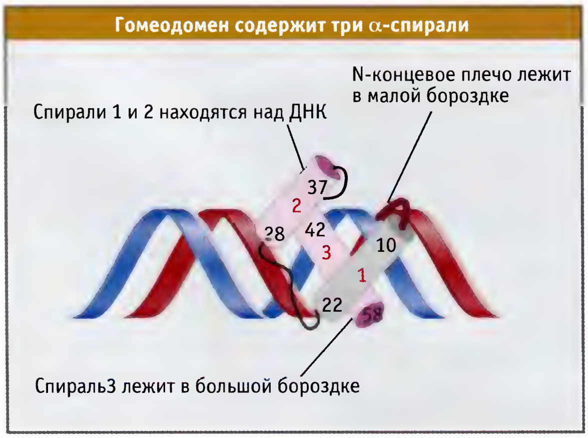


РИС. 25.25 Третья спираль гомеодомена укладывается в большую бороздку ДНК, а первая и вторая спирали гомеодомена располагаются за пределами двойной спирали ДНК. Таким образом, третья спираль контактирует и с фосфатным остовом двойной спирали, и со специфической последовательностью азотистых оснований. N-концевое плечо гомеодомена лежит в малой бороздке и образует дополнительные контакты с ДНК

Насколько универсальна эта модель, можно судить из сравнения кристаллической структуры гомеодоменов белка engrailed *D. melanogaster* и фактора спаривания α2 дрожжей *S. cerevisiae*. Строго говоря, ДНК-связывающий домен α2 не является гомеодоменом, но похож на него: те же три спирали, причем положение одной из этих спиралей можно почти точно совместить с аналогичной картинкой, полученной для engrailed. Скорее всего, все гомеодомены, даже не слишком похожие между собой, связывают ДНК в одной и той же манере. Это означает, что специфичность связывания во всех случаях задана аминокислотными остатками в немногочисленных конкретных положениях третьей α-спирали и N-концевого плеча гомеодомена.

Белки Hox, у которых есть гомеодомен, связывают ДНК с довольно низкой специфичностью к ее нуклеотидной последовательности. Долгое время вообще было непонятно, какова природа специфичности Hox как регуляторов. Как выяснилось, белки Hox входят в состав гетеродимеров; именно в таком виде они чаще всего связывают ДНК (партнерами им служат особые белки; у мух их называют Exd, у позвоночных животных — Pbx). Гетеродимеры (во всяком случае, *in vitro*) обладают более узкой специфичностью связывания ДНК, чем мономеры Hox; как правило, гетеродимер связывает 10 пн последовательность с консенсусом TGATNNATNN. Но даже этих знаний недостаточно для объяснения высокой специфичности регуляторов Hox. Есть еще один белок, Hth, необходимый для локализации Exd, партнера Hox, в клеточном ядре; он тоже входит в Hox-содержащий комплекс, который связывает ДНК, и, наверное, тоже способствует сужению специфичности связывания. Все равно непонятно, каким образом одни и те же партнеры, Exd и Hth, состоя в комплексе с произвольным членом семейства Hox, придают этому члену уникальную специфичность к конкретному набору мишеней; пока что этот вопрос остается открытым.

Белок, у которого есть гомеодомен, может быть либо активатором транскрипции, либо ее репрессором. Характер оказываемого эффекта зависит от других доменов; гомеодомен отвечает за одно лишь связывание с ДНК. Домены, придающие фактору транскрипции характер активатора либо репрессора, контролируются основным аппаратом транскрипции. Домены-активаторы могут контактировать с коактиваторами, которые, в свою очередь, уже должны иметь доступ к компонентам основного аппарата. Домены-репрессоры тоже должны, прямо или косвенно, взаимодействовать с компонентами аппарата транскрипции (то есть они никогда не перекрывают доступ к ДНК как таковой, принцип их действия не в этом). Например, у *Drosophila* репрессор Eve напрямую взаимодействует с $TF_{II}D$, одним из основных факторов транскрипции.

25.15 Белки с доменами типа «спираль-петля-спираль» димеризуются в различных сочетаниях

Ключевые понятия:

- Домен HLH («спираль-петля-спираль») имеет размер от 40 до 50 аминокислотных остатков; он состоит из двух амфипатических α-спиралей, размером по 15–16 остатков, разделенных петлей — неструктурированным сегментом полипептида.
- Амфипатические спирали ответственны за димеризацию.
- Белки bHLH имеют положительно заряженный мотив, примыкающий к домену HLH; этот мотив отвечает за связывание ДНК.
- В организме животного белки bHLH класса A экспрессируются повсеместно; белки bHLH класса B экспрессируются только в определенных тканях.
- Белки класса B образуют гетеродимеры с белками класса A.
- Белки HLH без положительно заряженного мотива (не bHLH), находясь в гетеродимерах с bHLH, мешают своим партнерам связывать ДНК.
- Белки HLH димеризуются в разных сочетаниях; соотношение этих сочетаний, за счет влияния на паттерн транскрипции, во многом определяет ход индивидуального развития организма.

ДНК-связывающие белки характеризуются наличием спиралевидных мотивов; кроме того, им свойственно димеризоваться. Обе эти особенности ярко выражены у группы белков, имеющих домены типа «спираль-петля-спираль», так называемых белков HLH (helix-loop-helix). Домены HLH имеют размеры от 40 до 50 аминокислотных остатков и содержат по две амфипатических α-спирали, разделенных неструктурированным промежутком, или петлей. Каждая амфипатическая спираль состоит из 15 или 16 аминокислотных остатков, среди которых есть несколько консервативных. По определению, амфипатическая спираль имеет две стороны: на одну сторону выходят гидрофобные (то есть электрически нейтральные) аминокислотные остатки, на другую сторону выходят гидрофильные (заряженные) остатки. Белки HLH образуют друг с другом гомо- и гетеродимеры. Таким образом, в каждой субъединице димера HLH есть две амфипатических спирали и связующая петля, которая дает им необходимую свободу: разнесенные в пространстве, спирали могут выстраивать паттерны своих контактов независимо друг от друга. Димеризация происходит в результате взаимодействия гидрофобных остатков на соответствующих сторонах двух амфипатических спиралей из разных субъединиц, так что склонность к димеризации, присущая всем белкам HLH, может быть полностью отнесена на счет амфипатических спиралей. На РИС. 25.26 приведен пример сравнения первичной структуры доменов HLH из разных белков.

У большинства белков HLH есть мотив, который, с химической точки зрения, можно квалифицировать как основание (то есть в нем содержится много положительно-заряженных аминокислотных остатков). Этот мотив расположен вплотную к одной из амфипатических спиралей домена HLH и нужен для связывания ДНК — это **bHLH-белки** (*basic* helix-loop-helix, что обозначает наличие мотива с основными химическими свойствами). При размере в ~15 аминокислотных остатков, ~6 из которых консервативны (см. рис. 25.26), положительно заряженный мотив позволяет димеру стабильно взаимодействовать с ДНК (при условии, что он есть у обеих субъединиц димера). Домены HLH явно играют роль организаторов, кооперирующих и ориентирующих ДНК-связывающие мотивы двух субъединиц димера.

Белки bHLH различаются по паттерну экспрессии. Некоторые из них экспрессируются повсеместно (однородно), то есть одинаково во всех тканях; Их объединяют в класс A; у млекопитающих к нему относятся факторы E12 и E47. Еще один класс белков bHLH, класс B, объединяет белки, встречающиеся лишь в отдельных тканях, то есть экспрессируемые с тканевой специфичностью. У млекопитающих к классу B относятся MyoD, миогенин и Myf-5 — в качестве активаторов транскрипции, все трое служат важнейшими регуляторами миогенеза (формирования мышечной ткани). Все белки bHLH класса B имеют один и тот же *modus operandi*: они димеризуются с белками класса A. Именно такие смешанные гетеродимеры осуществляют ключевую регуляцию различных дифференцировок. Например, у *D. melanogaster* развитие нервной системы находится под контролем *Ac-S* из класса B; партнером *Ac-S* при этом служит повсеместно

Белок HLH содержит два спиралевидных мотива

Белок	Последовательность	
MyoD	Ala Asp Arg Arg Lys Ala Ala Thr Met Arg Gln Arg Arg Arg	Положительно заряженный участок
Id	Arg Leu Pro Ala Leu Leu Asp Gln Glu Glu Val Asn Val Leu	В Id отсутствуют шесть консервативных аминокислот, придающих домену свойство основания
MyoD	Leu Ser Lys Val Asn Gln Ala Phe Gln Thr Leu Lys Arg Cys Thr	Спираль 1
Id	Leu Tyr Asp Met Asn Gly Cys Tyr Ser Arg Leu Lys Gln Leu Val	Консервативные аминокислотные остатки присутствуют в обоих белках, MyoD и Id
MyoD	Lys Val Gln Ile Leu Arg Asn Ala Ile Arg Tyr Ile Gln Gly Leu Glu	Спираль 2
Id	Lys Val Gln Ile Leu Glu His Val Ile Asp Tyr Ile Arg Asp Leu Glu	

РИС. 25.26 Все белки HLH содержат мотивы, соответствующие спиралям 1 и 2; их разделяет петля длиной от 10 до 24 аминокислот. Белки bHLH имеют участок, где сосредоточены консервативные положительно заряженные аминокислотные остатки; этот участок примыкает к спирали 1

экспрессируемый *da*. Белки Myc, продукты протоонкогенов и важнейшие регуляторы клеточного роста, выделяют в отдельный класс белков bHLH, со своим особым набором партнеров и мишеней.

Димеры белков bHLH различаются по своей способности к связыванию ДНК. Например, гомодимеры E47, гомодимеры E12–E47, а также гетеродимеры MyoD–E47 образуются легко и держатся на ДНК очень крепко. А вот, например, гомодимер E12, который тоже образуется легко, но плохо связывает ДНК. А MyoD (как и положено представителям класса B), вообще очень плохо гомодимеризуется. Соответственно, важными аспектами регуляции могут служить и образование димеров, и связывание димеров с ДНК. Чтобы описать этот тип регуляции, можно задать комбинаторное пространство, в котором представители разных классов bHLH попарно взаимодействуют друг с другом, причем все возможные пары потенциальных партнеров характеризуются своими вероятностями, эффективностями и, главное, последствиями своего образования для экспрессии генов-мишеней. К сожалению, эффективность димеризации и сродство димера к ДНК не могут быть предсказаны на основании одной лишь первичной структуры (аминокислотных последовательностей) потенциальных субъединиц.

Все димеры bHLH узнают в ДНК один и тот же консенсус; возможно, некоторые гомо- и гетеродимеры оказывают предпочтение конкретным вариантам мишеней; такие предпочтения, выказываемые димерами, могли бы углублять спецификацию их функции.

Сродство димеров к ДНК определяется свойствами отдельных сегментов полипептида внутри домена HLH (или в непосредственной близости от него). Например, в отличие от своего собрата E47, белок E12 содержит сегмент, служащий внутренним ингибитором, поскольку мешает соседнему (положительно заряженному) участку белковой глобулы взаимодействовать с ДНК. Именно поэтому гомодимеры E12, в отличие от гомодимеров E47, имеют такое низкое сродство к ДНК. У некоторых белков HLH положительные заряды, нацеленные на ДНК, гасятся пролинами либо вообще отсутствуют, как, например, у белка Id, показанного на рис. 25.26. Многие белки, похожие на Id (HLH, «bHLH без b»), охотно вступают в димеры с белками bHLH, но эти димеры совершенно неспособны к специфичному связыванию ДНК. Эта неспособность служит сильным подтверждением необходимости удвоения положительно заряженного мотива (мотива b в bHLH) для того, чтобы этот мотив мог выполнять функцию связывания ДНК.

Принципиальное различие между белками bHLH и неосновными белками HLH (у которых отсутствует положительно заряженный мотив) иллюстрируют свойства двух пар HLH-белков: *da-Ac-S/emc* и MyoD/Id-пары (РИС. 25.27). Они функционируют по принципу сетей регуляции.

У *D. melanogaster* есть ген *emc* (*extramacrochaetae*; у мутантов появляются «дополнительные большие щетинки»). Как следует из названия, ген отвечает за нормальное расположение органов чувств у взрослой мухи. Его продукт, белок emc, работает как *супрессор* эффекта работы нескольких генов, в числе которых гены *Ac-S* (полное название *achaete-scute*) и *da* (полное название *daughterless*). Гены *Ac-S* и *da* кодируют белки bHLH-типа. Их супрессор, ген *emc*, кодирует белок HLH, у которого отсутствует положительно заряженный мотив. По всей видимости, в отсутствие полноценного продукта *emc* белки *Ac-S* и *da* образуют димеры, активирующие транскрипцию определенных генов-мишеней в клетках, где (с точки зрения нормального развития) эта активация нежелательна, что и приводит к появлению фенотипических аномалий. В норме, продукция белка *emc* приводит к образованию гетеродимеров с его участием; такие гетеродимеры неспособны связывать ДНК. Продукция белка emc в определенных клетках необходима для подавления функций Ac/S.

Второй пример взят у совсем других животных. Диффереципровка мышечных клеток в индивидуальном развитии млекопитающих происходит как результат запрограммированного изменения профиля транскрипции. В реализации программы участвует несколько белков bHLH. Один из этих белков, MyoD, синтезируется только в миогенных клетках, которым суждено стать мышечными. (Речь идет о нормальной ситуации; в принципе, искусственное повышение экспрессии MyoD в некоторых других клетках может заставить эти клетки вступить на путь миогенной дифференцировки.) Будем считать, что спусковым механизмом миогенной дифференцировки служит гетеродимер MyoD/E12 (возможно, им служит гетеродимер MyoD/E47 и вряд ли им служит гомодимер MyoD). Перед началом миогенеза белок Id, у которого отсутствует положительно заряженный мотив, вступает в гетеродимеры с MyoD и/или E12 и/или E47. Такие димеры не способны связывать ДНК. Известно, что Id связывает E12 или E47 охотнее, чем MyoD. Создается впечатление, что до момента начала диффе-

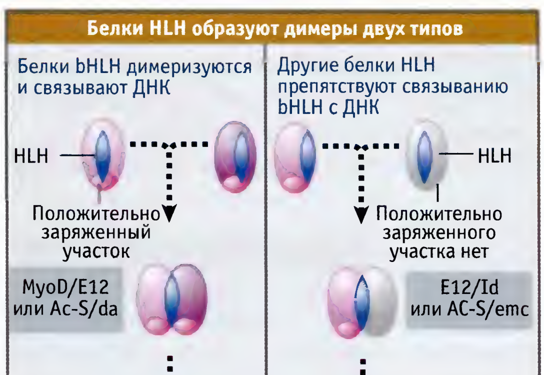


РИС. 25.27 Димер HLH, в котором обе субъединицы bHLH, может связывать ДНК. Димер HLH, в котором одна из субъединиц bHLH, а другая – нет, не может связывать ДНК

ренцировки Id забирает у MyoD (который есть только в закладках мышц) его потенциальных партнеров, E12 и E47 (которые есть во всех тканях развивающегося организма). Искусственное повышение экспрессии Id может полностью блокировать миогенез. Скорее всего, начало дифференцировки связано с уходом Id из сетки взаимодействий между белками HLH. После ухода Id белок MyoD, наконец, получает возможность проявить себя при поддержке E12 (или E47).

Активатор MyoD можно контролировать несколькими способами. Для предотвращения связывания с ДНК он может быть атакован HLH-партнером, таким как Id. В димере с E12 (или E47) MyoD эффективно активирует транскрипцию; оказывается, в димере с другим bHLH-белком, MyoR, тот же MyoD работает уже не как сайт-специфичеый активатор, но как сайт-специфичный репрессор. В пролиферирующих миобластах димер MyoD-MyoR подавляет транскрипцию в тех же самых локусах, для которых димер MyoD-E12/E47 служит активатором.

В целом организация белков HLH иллюстрирует два важнейших принципа регуляции транскрипции. Во-первых, сравнительно небольшое число белков-регуляторов выступает во всевозможных сочетаниях друг с другом, и у каждого сочетания может быть свой регуляторный смысл. Во-вторых, общая картина регуляции зависит от прихода и ухода конкретных неделимых компонентов регуляторных комплексов. В заключение отметим невероятную значимость регуляторов HLH: именно от них зависит спецификация многих клеточных типов, например, дифференциация сенсорных органов *Drosophila* или миогенная дифференцировка у человека.

25.16 «Лейциновые молнии» служат для димеризации

Ключевые понятия:

- «Лейциновая молния» объединяет две амфипатических спирали.
- Рядом с «молнией» находятся положительно заряженные участки субъединиц димера; эти участки контактируют с ДНК.
- Димеризация с использованием «лейциновой молнии» рождает домен bZIP; симметричные положительно заряженные участки этого домена связывают инвертированные повторы в ДНК.

В построении комплекса транскрипции взаимодействия между белками играют одну общую тему. Аминокислотные мотивы, участвующие в этих взаимодействиях, довольно однообразны. В событиях регуляции, между собой никак не связанных, могут быть задействованы одни и те же (инвариантные) мотивы, которые можно встретить в различных активаторах, а также в других белках. «Лейциновая молния» представляет собой тип взаимодействия, свойственный многим белкам-регуляторам. Для него характерен очень выразительный паттерн белок-белковых контактов. «Лейциновая молния» объединяет в себе два мотива из разных полипептидов (недаром ее сравнивают с застежкой из двух зацепленных друг за друга половинок). Как известно, образование димеров свойственно белкам, узнающим в ДНК определенные последовательности-мишени. В случае с «лейциновыми молниями» смысл димеризации особенно прозрачен: мы буквально видим, как димеризация подравнивает ДНК-связывающие мотивы двух субъединиц, придает четкость их положениям друг относительно друга, фиксирует расположение всего димера относительно ДНК. Связывание димера с ДНК схематично изображено на РИС. 25.28.

Амфипатическая α-спираль представляет собой особый мотив, вычленяемый у многих активаторов. В амфипатической спирали гидрофобные группы аминокислотных остатков (в том числе лейциновых) смотрят в одну сторону; гидрофильные (то есть заряженные, полярные) группы смотрят в другую сторону. «Лейциновую молнию» можно квалифицировать как гибридную α-спираль, в которой лейцины, выступающие из амфипатической α-спирали одной субъединицы, переплетены с лейцинами из такой же спирали другой субъединицы; вместе они образуют спираль второго порядка — суперспираль (coiled coil). На виток каждой из двух (закрученных друг вокруг друга) правосторонних α-спиралей приходится по три с половиной аминокислотных остатка, так что весь паттерн вращения воспроизводится с интервалом в семь аминокислотных остатков.

Какое отношение к связыванию ДНК может иметь такая структура? Дело в том, что в каждой субъединице димера мотив из повторов, перемежаемых лейцинами (то есть половинка «молнии») соседствует с положительно заряженным участком, похожим на ДНК-связывающий мотив. Застегнутая «лейциновая молния» дает Y-образную структуру. (Ножка Y — это сама «молния»; от нее симметрично отходят положительно заряженные мотивы, связывающие ДНК.) Ее обозначают как **bZIP-мотив**. Мишень bZIP (то есть сайт связывания димера, «застегнутого на молнию») представляет собой слитные



РИС. 25.28 Положительно заряженные участки доменов bZIP удерживаются рядом за счет гидрофобных взаимодействий: спиралевидные мотивы двух субъединиц взаимодействуют с образованием «лейциновой молнии» и димеризуются в параллельной ориентации

инвертированные повторы, поскольку в «застегнутом» димере ДНК-связывающие домены двух субъединиц практически сомкнуты.

«Лейциновые молнии» могут использоваться для образования как гомо-, так и гетеродимеров. Каждая половинка «молнии», довольно протяженный мотив, состоит из приблизительных повторов аминокслотной последовательности. В каждом седьмом положении повторы перемежаются лейцинами или другими гидрофобными остатками. Так, в белке C/EBP насчитывается четыре повтора Leu-X_6, а в белках Jun и Fos насчитывается по пять таких повторов. (Димер C/EBP связывает CAAT-бокс, а также энхансер SV40; гетеродимер Jun и Fos известен как активатор AP1.)

Активатор AP1 был первоначально идентифицирован по специфичному связыванию энхансера SV40 (см. рис. 24.24). Активный препарат AP1 содержит несколько различных полипептидов. Самый обильный из них, Jun, является продуктом гена *c-jun*, родственного гену *v-jun*, переносимому птичьим вирусом саркомы. В геноме мыши есть целое семейство генов *jun*; кроме первично идентифицированного *c-jun* в это семейство входят гены *junB* и *junD*, выявленные на основании гомологии с *jun*. Как продукты родственных генов все три белка Jun похожи между собой; они образуют друг с другом гомо-и гетеродимеры, соединенные «лейциновыми молниями».

Другой обильный компонент гетерогенного фактора AP1, белок Fos, является продуктом гена *c-fos*; у него тоже есть онкогенный гомолог, ген *v-fos*, переносимый мышиным вирусом саркомы. Экспрессия гена *c-fos* приводит к активации транскрипции генов, в промоторах которых есть сайты-мишени AP1. Белок Fos (фосфорилированный белок, локализованный в ядре) представляет целое семейство родственных факторов. Остальные представители семейства известны как антигены, родственные Fos (FRA, Fos-related antigens; или Fos-подобные белки — Fos-like proteins).

Хотя у Fos тоже есть потенциал к образованию «лейциновой молнии», его субъединицы не димеризуются между собой; зато они легко образуют «лейциновые молнии» с Jun. Способность Jun и Fos к образованию гетеродимеров принципиально важна для взаимодействия этих факторов с ДНК. Если нет Jun, Fos не может связывать ДНК (наверное, оттого, что не образует гомодимеров). Находясь в составе гетеродимеров с Jun, тот же самый Fos прекрасно связывает ДНК. Хотя гетеродимеры Jun/Fos и гомодимеры Jun/Jun связывают сайты AP1 одинаково специфично, гетеродимеры связывают их в десять раз крепче.

25.17 Резюме

К факторам транскрипции относятся основные факторы, активаторы и коактиваторы. Основные факторы взаимодействуют с РНК-полимеразой в точке старта транскрипции. Активаторы специфично взаимодействуют с короткими участками ДНК — элементами ответа, находящимися в промоторах или энхансерах. Активаторы проявляют себя через белок-белковые взаимодействия с компонентами основного аппарата. Некоторые активаторы взаимодействуют с основным аппаратом напрямую; другие нуждаются в посредничестве коактиваторов. Многие активаторы построены из модулей. В таких активаторах можно вычленить два независимых домена, один из которых отвечает за связывание ДНК; другой домен отвечает за саму активацию. ДНК-связывающий домен нужен для удержания активирующего домена в пределах досягаемости белкового комплекса инициации транскрипции.

Элементы ответа бывают разными. Некоторые из них есть у многих генов, а работают с ними факторы, которые есть во всех клетках организма. Некоторые элементы ответа есть лишь у ограниченного числа генов и узнаются факторами, специфичными для конкретных клеточных типов.

Промоторы РНК-полимеразы II содержат огромное многообразие *цис*-действующих коротких элементов, с которыми, в *транс*-манере, работают факторы транскрипции. В промоторе такие элементы предшествуют TATA-боксу; ориентация элементов и расстояние от них до стартовой точки могут быть любыми. От взаимодействий активаторов со своими элементами зависит эффективность использования промотора в конкретный момент времени. Каждый активатор имеет в основном аппарате транскрипции свою мишень для установления белок-белкового контакта (или непосредственно, или через коактиватор; именно этот контакт позволяет активатору выполнять активацию). Мишенями различным активаторам служат такие компоненты основного аппарата, как $TF_{II}D$, $TF_{II}B$, $TF_{II}A$. Взаимодействие активатора с TF стимулирует сборку основного аппарата.

Многие активаторы обладают серией характерных аминокислотных мотивов, известных как «цинковые пальцы» и служащих для связывания ДНК. В каждом «цинковом пальце» есть остатки цистеина, которые удерживают атом цинка. «Цинковые пальцы» бывают двух типов: Cys_2/His_2 «пальцы» обязательно многочисленны, в то время как Cys_2/Cys_2 «палец» у белка может быть и единственным (либо их два).

Отдельные классы факторов транскрипции были идентифицированы на основании сравнения гомологичных нуклеотидных последовательностей генов, кодирующих эти факторы. Один из таких классов — это гомеодомен-содержащие белки; к ним относятся важнейшие регуляторы раннего развития у насекомых и круглых червей, а также многие факторы транскрипции млекопитающих. Гомеодомен состоит из ~60 аминокислотных остатков; у него есть явное сродство с прокариотическими ДНК-связывающими доменами типа «спираль-поворот-спираль». Гомеодомен служит факторам для связывания ДНК.

Мотив «лейциновая молния» богат лейцинами и служит факторам для димеризации. В каждой субъединице димера к этому мотиву примыкает положительно заряженный мотив, служащий фактору для связывания ДНК.

Рецепторы стероидных гормонов были идентифицированы как первые представители факторов транскрип-

ции, включаемых (активируемых) за счет связывания небольшого гидрофобного гормона. Активированный фактор локализуется в ядре; опираясь на «свои» элементы в регуляторных областях некоторых генов, фактор активирует транскрипцию этих генов. Факторы этого класса имеют ДНК-связывающие домены с «цинковыми пальцами»; они связывают ДНК в составе гомо- или гетеродимеров. Все гомодимеры узнают пары инвертированных повторов с общим консенсусом; элементом специфичности служит интервал между повторами. Гетеродимеры узнают пары прямых повторов; опять же, разные типы гетеродимеров узнают разный размер интервала, разделяющего повторы. У рецепторов стероидных гормонов ДНК-связывающие домены имеют по два «цинковых пальца». Первый «палец» задает специфичность к консенсусу, второй «палец» реагирует на определенный интервал между повторами (половинками узнаваемого элемента).

Белки с доменами HLH (helix-loop-helix; «спираль-петля-спираль») содержат амфипатические α-спирали, ответственные за димеризацию. У большинства белков HLH ответственные за димеризацию амфипатические α-спирали соседствуют с положительно заряженными мотивами, связывающими ДНК. Такие белки называют bHLH (basic helix-loop-helix); по паттерну присутствия их принято делить на повсеместно (однородно) экспрессируемые и тканеспецифичные. Типичный активный белок bHLH представляет собой гетеродимер из двух субъединиц; одна из них встречается во всех клетках, другая — только в клетках определенного типа. Димер, у которого хотя бы одна из субъединиц не имеет положительно заряженного участка, неспособен связывать ДНК; поэтому такие «холостые» субъединицы работают как отрицательные регуляторы экспрессии генов. От комбинаторной ассоциации различных субъединиц образуются разветвленные регуляторные сети.

Очень многие факторы транскрипции работают именно в форме димеров; семейства таких факторов устроены как наборы различных гомо- и гетеродимеров. Это придает регуляции комбинаторную мощность; при этом экспрессию генов можно рассматривать как функцию от набора сочетаний факторов (изменяемого во времени и пространстве). К некоторым классам активаторов транскрипции приходится также отнести и гомологичные репрессоры транскрипции. В некоторых случаях вклад такого репрессора в образование димера с гомологичным партнером-активатором равносилен ингибированию функции этого партнера. В альтернативных случаях репрессия идет через контакт репрессора с корепрессором (подобно активации, идущей через контакт активатора с коактиватором).

Литература

25.2 Есть несколько разновидностей факторов транскрипции

Обзор

Lee, T. I. and Young, R. A. (2000). Transcription of eukaryotic protein-coding genes. *Annu. Rev. Genet.* 34, 77–137.

Lemon, B. and Tjian, R. (2000). Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev.* 14, 2551–2569.

25.3 Связывание ДНК и активация транскрипции осуществляются независимыми доменами фактора

Обзор

Guarente, L. (1987). Regulatory proteins in yeast. *Annu. Rev. Genet.* 21, 425–452.

Ptashne, M. (1988). How eukaryotic transcriptional activators work. *Nature* 335, 683–689.

25.4 Двугибридная система была разработана для поиска новых белок-белковых взаимодействий

Исследование

Fields, S. and Song, O. (1989). A novel genetic system to detect protein-protein interactions. *Nature* 340, 245–246.

25.5 Активаторы взаимодействуют с основным аппаратом транскрипции

Обзор

Lemon, B. and Tjian, R. (2000). Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev.* 14, 2551–2569.

Maniatis, T., Goodbourn, S., and Fischer, J. A. (1987). Regulation of inducible and tissue-specific gene expression. *Science* 236, 1237–1245.

Mitchell, P., and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA-binding proteins. *Science* 245, 371–378.

Myers, L. C. and Kornberg, R. D. (2000). Mediator of transcriptional regulation. *Annu. Rev. Biochem.* 69, 729–749.

Исследование

Asturias, F. J., Jiang, Y. W., Myers, L. C., Gustafsson, C. M., and Kornberg, R. D. (1999). Conserved structures of mediator and PHK polymerase II holoenzyme. *Science* 283, 985–987.

Chen, J.-L. et al. (1994). Assembly of recombinant $TF_{II}D$ reveals differential coactivator requirements for distinct transcriptional activators. *Cell* 79, 93–105.

Dotson, M. R., Yuan, C. X., Roeder, R. G., Myers, L. C., Gustafsson, C. M., Jiang, Y. W., Li, Y., Kornberg, R. D., and Asturias, F. J. (2000). Structural organization of yeast and mammalian mediator complexes. *Proc. Natl. Acad. Sci. USA* 97, 14307–14310.

Dynlacht, B. D., Hoey, T., and Tjian, R. (1991). Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell* 66, 563–576.

Kim, Y. J., Bjorklund, S., Li, Y., Sayre, M. H., and Kornberg, R. D. (1994). A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of PHK polymerase II. *Cell* 77, 599–608.

Ma, J. and Ptashne, M. (1987). A new class of yeast transcriptional activators *Cell* 51, 113–119.

Pugh, B. F. and Tjian, R. (1990). Mechanism of transcriptional activation by Sp1: evidence for coactivators. *Cell* 61, 1187–1197.

25.6 Некоторые связывающие промотор белки являются репрессорами

Исследование

Goppelt, A., Stelzer, G., Lottspeich, F., and Meisterernst, M. (1996). A mechanism for repression of class II gene transcription through specific binding of NC2 to TBP-promoter complexes via heterodimeric histone fold domains. *EMBO J.* 15, 3105–3116.

Inostroza, J. A., Mermelstein, F. H., Ha, I., Lane, W. S., and Reinberg, D. (1992). Dr1, a TATA-binding protein-associated phosphoprotein and inhibitor of class II gene transcription. *Cell* 70, 477–489.

Kim, T. K., Zhao, Y., Ge, H., Bernstein, R., and Roeder, R. G. (1995). TATA-binding protein residues implicated in a functional interplay between negative cofactor NC2 (Dr1) and general factors $TF_{II}A$ and $TF_{II}B$. *J. Biol. Chem.* 270, 10976–10981.

25.8 ДНК-связывающие домены бывают разными

Обзор

Harrison, S. C. (1991). A structural taxonomy of DNA-binding proteins. *Nature* 353, 715–719.

Pabo, C. T. and Sauer, R. T. (1992). Transcription factors: structural families and principles of DNA recognition. *Annu. Rev. Biochem.* 61, 1053–1095.

Исследование

Miller, J. et al. (1985). Repetitive zinc binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J.* 4, 1609–1614.

Murre, C., McCaw, P. S., and Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* 56, 777–783.

25.9 «Цинковые пальцы» — это ДНК-связывающие домены

Исследование

Kadonaga, J. et al. (1987). Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. *Cell* 51, 1079–1090.

Miller, J. et al. (1985). Repetitive zinc binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J.* 4, 1609–1614.

Pavletich, N. P. and Pabo, C. O. (1991). Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 21 Å. *Science* 252, 809–817.

25.10 Рецепторы стероидов являются активаторами

Обзор

Evans, R. M. (1988). The steroid and thyroid hormone receptor superfamily. *Science* 240, 889–895.

Mangelsdorf, D. J. and Evans, R. (1995). The RXR heterodimers and orphan receptors. *Cell* 83, 841–850.

25.11 Рецепторы стероидов имеют «цинковые пальцы»

Обзор

Tsai, M J, and O'Malley, B. W. (1994). Molecular mechanisms of action of steroid/thyroid receptor superfamily members. *Annu. Rev. Biochem.* 63, 451–486.

Исследование

Umesono, K. and Evans, R. M. (1989). Determinants of target gene specificity for steroid/thyroid hormone receptors. *Cell* 57, 1139–1146.

25.13 Рецепторы стероидов узнают элементы ответа в соответствии с комбинаторным кодом

Обзор

Mangelsdorf, D. J. and Evans, R. (1995). The RXR heterodimers and orphan receptors. *Cell* 83, 841–850.

Yamamoto, K. R. (1985). Steroid receptor regulated transcription of specific genes and gene networks. *Annu. Rev. Genet.* 19, 209–252.

Исследование

Hurlein, A. J. et al. (1995). Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor corepressor. *Nature* 377, 397–404.

Rastinejad, F., Perlmann, T., Evans, R. M., and Sigler, P. B. (1995). Structural determinants of nuclear receptor assembly on DNA direct repeats. *Nature* 375, 203–211.

Umesono, K., Murakami, K. K., Thompson, C. C., and Evans, R. M. (1991). Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors. *Cell* 65, 1255–1266.

25.14 Гомеодомены узнают в ДНК стандартную последовательность-мишень

Обзор

Gehring, W. J. et al. (1994). Homeodomain-DNA recognition. *Cell* 78, 211–223.

Исследование

Han, K., Levine, M. S., and Manley, J. L. (1989). Synergistic activation and repression of transcription by *Drosophila* homeobox proteins. *Cell* 56, 573–583.

Wolberger, C. et al. (1991). Crystal structure of a MATα2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell* 67, 517–528.

25.15 Белки с доменами типа «спираль-петля-спираль» димеризуются в различных сочетаниях

Обзор

Weintraub, H. (1991). The MyoD gene family: nodal point during specification of the muscle cell lineage. *Science* 251, 761–766.

Исследование

Benezra, R. et al. (1990). The protein Id: a negative regulator of helix-loop-helix DNA-binding proteins. *Cell* 61, 49–59.

Davis, R. L. et al. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* 51, 987–1000.

Davis, R. L. et al. (1990). The MyoD DNA binding domain contains a recognition code for muscle-specific gene activation. *Cell* 60, 733–746.

Lassar, A. B. et al. (1991). Functional activity of myogenic HLH proteins requires hetero-oligomerization with E12/E47-like proteins *in vitro*. *Cell* 66, 305–315.

Murre, C., McCaw, P. S., and Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* 56, 777–783.

25.16 «Лейциновые молнии» служат для димеризации

Обзор

Vinson, C. R., Sigler, P. B., and McKnight, S. L. (1989). Scissors-grip model for DNA recognition by a family of leucine zipper proteins. *Science* 246, 911–916.

Исследование

Landschulz, W. H., Johnson, P. F., and McKnight, S. L. (1988). The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* 240, 1759–1764.

Сплайсинг и процессинг РНК

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

26.1 Введение

26.2 Сайты сплайсинга ядерных генов представляют собой короткие последовательности

- Сайты сплайсинга — последовательности, расположенные на границах экзон-интрон. Их называют согласно расположению в интроне.
- 5′-граница сплайсинга, расположенная на 5′(левом)-конце интрона, содержит консенсусную последовательность GU.
- 3′-граница сплайсинга, расположенная на 3′(правом)-конце интрона, содержит консенсусную последовательность AG.
- Согласно правилу GU-AG (изначально в контексте ДНК, называемому правилом GT-AG), эти динуклеотиды должны располагаться в двух первых и двух последних позициях всех интронов в последовательности пре-мРНК.

26.3 Границы сплайсинга распознаются аппаратом сплайсинга в парах

- Сплайсинг зависит только от распознавания пар сайтов сплайсинга.
- Все 5′-сайты функционально эквивалентны между собой, равно как неразличимы и все 3′-сайты.

26.4 В ходе сплайсинга мРНК образуется структура типа лассо

- В сплайсинге участвуют три сайта: 5′- и 3′-сайты сплайсинга, а также сайт разветвления, расположенный перед 3′-сайтом сплайсинга.
- Последовательность сайта разветвления высококонсервативна у дрожжей и менее консервативна у высших эукариот.
- После разрезания в 5′-сайте сплайсинга образовавшийся 5′-конец интрона взаимодействует с 2′-положением азотистого основания А, расположенного в сайте разветвления. При этом формируется структура типа лассо.
- После разрезания в 3′-сайте спласинга интрон высвобождается в форме лассо, а левый и правый экзоны лигируются.
- Сплайсинг реализуется посредством реакций трансэтерификации, в которых связь переносится из одного положения в другое.

26.5 мяРНК участвуют в сплайсинге

- К пяти мяРНК, вовлеченным в процесс сплайсинга, относятся U1, U2, U5, U4 и U6.
- Вместе с несколькими дополнительными белками мяРНК формируют сплайсосому.
- За исключением U6, все мяРНК содержат консервативную последовательность, связывающую Sm-белки, распознаваемые антителами, образующимися при аутоиммунном заболевании.

26.6 U1 мяРНП инициирует сплайсинг

- U1 мяРНП инициирует сплайсинг, связываясь с 5′-сайтом сплайсинга с помощью РНК-РНК взаимодействий.
- В состав комплекса E входит U1 мяРНП, связанная с 5′-сайтом сплайсинга, белок U2AF, связанный с пиримидиновым регионом между сайтом разветвления и 3′-сайтом сплайсинга, и группа SR-белков, соединяющих U1 мяРНП с U2AF.

26.7 Формирование комплекса E может происходить в результате распознавания как интронов, так и экзонов

- Прямой путь формирования комплекса E представляет собой связывание U1 мяРНП с 5′-сайтом сплайсинга, и U2AF — с пиримидиновым регионом, расположенным между сайтом разветвления и 3′-сайтом сплайсинга.
- Другая возможность заключается во взаимодействии U2AF с пиримидиновым регионом и U1 мяРНП с 5′-сайтом сплайсинга следующего интрона.
- Комплекс E превращается в комплекс A в результате связывания U2 мяРНП с сайтом разветвления.
- В случае когда комплекс E формируется при участии 5′-сайта сплайсинга следующего интрона, при превращении комплекса E в комплекс A этот сайт сплайсинга заменяется правильным 5′-сайтом.
- Слабый 3′-сайт сплайсинга требует участия энхансеров сплайсинга – последовательностей, расположенных в правом экзоне и способных напрямую связываться с SR-белками.

26.8 Пять мяРНП формируют сплайсосому

- При связывании U5 и U4/U6 мяРНП комплекс A превращается в сплайсосому B1, которая уже содержит все компоненты, необходимые для сплайсинга.
- Во время сплайсинга сплайсосома подвергается изменениям, последовательно проходя несколько стадий.
- Высвобождение U1 мяРНП дает возможность U6 мяРНП взаимодействовать с 5′-сайтом сплайсинга и обеспечивает превращение сплайсосомы B1 в сплайсосому B2.

- Диссоциация U4 мяРНП от U6 мяРНП приводит к спариванию U6 с U2 и формированию каталитически активного сайта.

26.9 Аппарат альтернативного сплайсинга использует различные мяРНП

- В реакции альтернативного сплайсинга задействован другой набор мяРНП. Комплекс, который они формируют, называют сплайсосомой U12.
- Интроны-мишени можно идентифицировать с помощью длинных консенсусных последовательностей на границах сплайсинга, однако обычно они содержат те же самые GU-AG динуклеотиды.
- В некоторых интронах на границах сплайсинга расположены динуклеотиды AU-AC. Часть из этих интронов U1-зависимые, часть — U12-зависимые.

26.10 Сплайсинг сопряжен с экспортом мРНК

- Белки REF связывают сайты сплайсинга, взаимодействуя со сплайсосомой.
- После окончания сплайсинга они остаются связанными с РНК в области границы экзон-экзон.
- Белки REF взаимодействуют с транспортным белком TAP/Mex, экспортирующим РНК сквозь ядерную пору.

26.11 Автосплайсинг интронов группы II происходит через формирование лассо

- Интроны группы II, обладая автокаталитической активностью, способны к самовырезанию из РНК.
- Границы сплайсинга и механизмы сплайсинга интронов группы II аналогичны таковым для ядерных интронов.
- Вторичная структура интрона группы II формирует каталитический сайт, напоминающий структуру ядерного интрона U6-U2.

26.12 Альтернативный сплайсинг — это реализация различных вариантов сшивания экзонов между собой

- Отдельные экзоны могут исключаться из РНК продукта или, наоборот, включаться в него в результате сшивания между собой разных пар сайтов сплайсинга.
- При использовании альтернативных вариантов сшивания экзонов длина конечного продукта сплайсинга может меняться.
- Детерминация пола у *Drosophila* является результатом серии реакций альтернативного сплайсинга в генах, кодирующих последовательные компоненты данного процесса.
- Р-элементы *Drosophila* подвергаются зародыш-специфичному альтернативному сплайсингу.

26.13 В реакциях *транс*-сплайсинга участвуют малые РНК

- Реакции сплайсинга, как правило, являются *цис*-реакциями, то есть происходят между участками одной молекулы РНК.
- *Транс*-сплайсинг, свойственный трипаносомам и червям, происходит, когда короткая последовательность (СЛ РНК) сплайсируется с 5'-концами разных предшественников мРНК.
- СЛ РНК имеет структуру, напоминающую структуру Sm-связывающего сайта U мяРНК, и, возможно, играет в реакции аналогичную роль.

26.14 Сплайсинг тРНК у дрожжей идет через разрезание и сшивание предшественника

- Сплайсинг тРНК представляет собой серию реакций разрезания и лигирования.

26.15 Эндонуклеаза сплайсинга распознает тРНК

- Эндонуклеаза расщепляет предшественник тРНК по обоим концам интрона.
- Эндонуклеаза дрожжей представляет собой гетеротетрамер с двумя (родственными) каталитическими субъединицами.
- Для распознавания сайтов расщепления эндонуклеаза использует механизмы, позволяющие оценить положение фермента относительно определенного участка тРНК.
- Нуклеаза архей имеет более простую структуру и распознает мотив «изгиб-спираль-изгиб» в субстрате.

26.16 Расщепление и лигирование тРНК — две отдельные реакции

- В результате вырезания интрона образуются две половины молекулы тРНК, которые, соединяясь друг с другом, формируют зрелую структуру.
- Эти половины имеют необычные концы: на 5'-конце — гидроксильная группа, а на 3'-конце — 2'–3'-циклический фосфат.
- 5'-OH-группа фосфорилируется полинуклеотидкиназой, циклическая фосфатная группа под действием фосфодиэстеразы раскрывается и образует 2'-фосфатный конец и 3'-OH-группу, концы экзонов соединяются РНК-лигазой, а 2'-фосфат удаляется фосфатазой.

26.17 Каскад реакций, запускаемых в ответ на накопление белков, связан с процессом сплайсинга тРНК

- Ire1p – белок внутренней ядерной мембраны, N-концевой домен этого белка обращен в просвет ЭПР, а С-концевой домен – внутрь ядра.
- Связывание несвернувшегося белка с N-терминальным доменом запускает реакцию автофосфорилирования, в результате чего активируется С-терминальная нуклеаза.
- Активированная нуклеаза расщепляет мРНК Hac1, при этом интрон высвобождается, а экзоны лигируются тРНК-лигазой.
- Сплайсированная мРНК Hac1 кодирует транскрипционный фактор, активирующий гены шаперонов, способствующих фолдингу белков.

26.18 3'-концы транскриптов polI и polIII образуются в результате терминации транскрипции

- РНК-полимераза I терминирует транскрипцию, достигая 18-нуклеотидной терминаторной последовательности.
- РНК-полимераза III терминирует транскрипцию, достигая поли$(U)_4$ последовательности, расположенной внутри G-C богатого региона.

26.19 3'-концы мРНК образуются в результате расщепления и полиаденилирования первичного транскрипта

- Последовательность AAUAAA является маркером 3'-конца, образующегося в результате расщепления молекулы мРНК с последующим полиаденилированием.
- Реакция происходит при участии белкового комплекса, в состав которого входят фактор специфичности, эндонуклеаза и поли(А)-полимераза.
- Фактор специфичности и эндонуклеаза расщепляют РНК в участке, расположенном после последовательности AAUAAA.

- Фактор специфичности и поли(А)-полимераза добавляют примерно 200 остатков А к формируемому 3′-концу.
- A-U-богатая последовательность на 3′-конце мРНК контролирует процессы поли- и деаденилирования в цитоплазме эмбрионов *Xenopus*.

- мРНК гистонов не полиаденилированы; их 3′-конец образуется в результате реакции расщепления, зависящей от структуры мРНК.
- Условием осуществления реакции расщепления является связывание SLBP с структурой типа стебель-петля на мРНК и спаривание U7 мяРНК с прилегающим к этой структуре одноцепочечным регионом.

- Большие и малые рРНК образуются в результате расщепления общего предшественника.

- C/D-группа snoРНК необходима для метилирования рибозы в 2′-положении.
- H/ACA-группа snoРНК необходима для превращения уридина в псевдоуридин.
- В каждом случае snoРНК спаривается с последовательностью рРНК, содержащей основания-мишени. При этом формируется структура, являющаяся субстратом для модификации.

26.1 Введение

Прерывистые гены обнаружены во всех классах живых организмов. В геномах низших эукариот они представлены в небольшом количестве, в то время как у более высокоорганизованных эукариот на их долю приходится бо́льшая часть генома. Отдельные гены различаются числом и длиной интронов, тем не менее типичный ген млекопитающего состоит из 7–8 экзонов и имеет длину ~16 тпн. Длина экзонов составляет порядка 100–200 пн, а интронов — более 1 тпн (см. разд. 3.6 «Длина генов сильно варьирует»).

Несоответствие между прерывистой организацией гена и непрерывной структурой мРНК свидетельствует о том, что первичный транскрипт подвергается процессингу. Первичный транскрипт повторяет организацию гена и называется **пре-мРНК**. Типичная молекула мРНК, образующаяся при удалении интронов из пре-мРНК, имеет длину примерно 2,2 тпн. Процесс удаления интронов называют **сплайсингом РНК**. Вырезание интронов — важнейший этап формирования зрелых молекул мРНК у всех эукариот. (Тот факт, что прерывистые гены у низших эукариот, таких как дрожжи, встречаются сравнительно редко, не отражает в полной мере значимости интронов. Основные белки высших эукариот кодируются именно прерывистыми генами, а большая часть суммарной мРНК клетки (возможно, более 50%, что не очевидно из анализа генома) образуется при задействовании механизмов сплайсинга.)

Свойства ядерной РНК высших эукариот отражают природу несоответствия размеров ядерных генов размерам их продуктов. Средний размер и сложность нуклеотидной последовательности ядерной РНК существенно превосходят размер и сложность состава молекул мРНК. Поскольку размеры ядерной РНК колеблются в очень широких пределах, ее называют **гетерогенной ядерной РНК (гяРНК).** В основном пул гяРНК представлен пре-мРНК, однако может включать и другие транскрипты (транскрипты, которые не дозревают с образованием мРНК).

гяРНК входит в состав рибонуклеопротеиновой частицы (**гяРНП**), в которой гяРНК связана с белками. Как изображено на РИС. 26.1, структура гяРНП, описанная *in vitro,* представляет собой сферические элементы, соединенные фибриллами. Белки, представленные в гяРНП в наибольшем количестве, составляют ядро частицы. Помимо основных белков, в состав гяРНП входит порядка 20 белков, присутствующих в небольшом количестве. Как правило, примерно на 10^8 копий молекул белка в ядре приходится около 10^6 молекул гяРНК. Некоторые из этих белков могут играть структурную роль в упаковке молекул гяРНК, участвовать в транспортировке РНК между ядром и цитоплазмой или контролировать активность РНК иным способом.

Как и все другие модификации, которым подвергается вновь синтезированная РНК, реакции сплайсинга протекают в ядре. На РИС. 26.2 перечислены этапы экспрессии прерывистых генов. Происходит кэпирование первичного транскрипта с 5′-конца (см. разд. 7.9 «У эукариот 5′-конец мРНК кэпирован»), полиаденилирование с 3′-конца (см. разд. 7.10 «3′-конец полиаденилирован») и вырезание из его состава интронов. После этого РНК

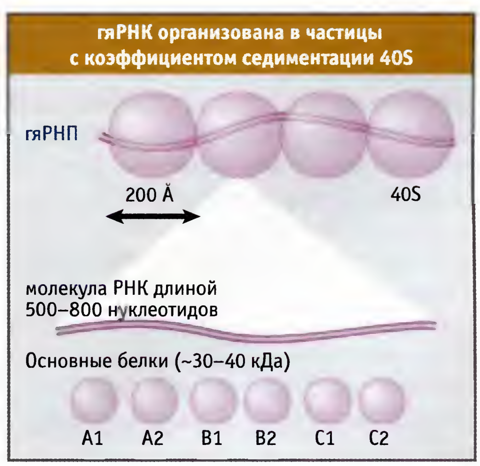


РИС. 26.1 гяРНК существует в виде рибонуклеопротеиновых частиц, организованных в серию сферических структур

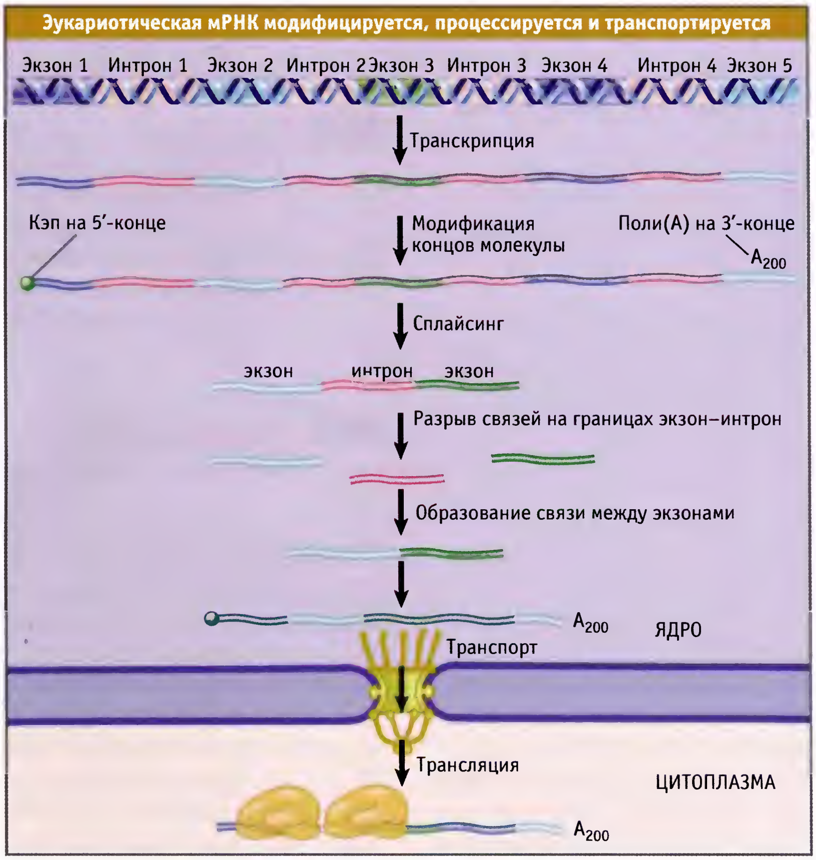


РИС. 26.2 Реакции модификации РНК в ядре включают в себя формирование 5′- и 3′-концов и удаление интронов в результате сплайсинга. Сплайсинг заключается в разрыве связей на границах экзон–интрон и образовании связи между концами экзонов. Молекулы зрелой мРНК транспортируются сквозь ядерные поры в цитоплазму, где вовлекаются в реакции трансляции

транспортируется через ядерные поры в цитоплазму, где вовлекается в процессы трансляции.

Рассматривая сплайсинг РНК как один из механизмов экспрессии генов, интересно выяснить, какие участки РНК затрагивает этот процесс (по сравнению с другими модификациями РНК). Происходит ли он в определенном участке ядра, связан ли с другими процессами, такими, как, например, транспорт из ядра в цитоплазму? Могут ли нарушения в механизмах сплайсинга привести к существенным различиям в экспрессии прерывистых и непрерывных генов?

Главным вопросом в изучении природы реакций сплайсинга является контроль его специфичности. Какие механизмы лежат в основе точного распознавания концов интронов и вырезания из РНК строго определенных последовательностей? Удаляются ли интроны из предшественника в определенном порядке? Является ли процесс созревания РНК фактором *регуляции* экспрессии генов, осуществляющейся путем отбора из имеющихся предшественников или путем изменения характера сплайсинга?

Мы можем выделить несколько типов систем сплайсинга:

- Удаление интронов из ядерных РНК высших эукариот осуществляется системой, которая распознает короткие консенсусные последовательности, расположенные на границах экзон–интрон и внутри интронов. Осуществление реакций сплайсинга по этому принципу требует участия большого аппарата, состоящего из ряда белков и нуклеопротеинов, функционирующих в составе комплекса, называемого сплайсосомой. Данный механизм сплайсинга реализуется посредством реакции трансэтерификации, которая катализируется как белками, так и РНК.
- Некоторые РНК обладают способностью автономно осуществлять вырезание интронов из своей структуры. Интроны такого типа делят на две группы, отличающиеся друг от друга вторичной и третичной структурами. В обеих группах сплайсинг включает в себя реакцию трансэтерификации, в которой в роли катализирующего агента выступает РНК (см. гл. 27 «Каталитическая РНК»).
- Удаление интронов из ядерных предшественников тРНК дрожжей происходит при участии ферментов, взаимодействие которых с субстратом напоминает действие ферментов процессинга тРНК, поскольку решающим фактором в обоих случаях служит конформация предшественника тРНК. Реакции сплайсинга включают в себя расщепление и лигирование субстрата.

26.2 Сайты сплайсинга ядерных генов представляют собой короткие последовательности

Ключевые понятия:

- Сайты сплайсинга — последовательности, расположенные на границах экзон-интрон. Их называют согласно расположению в интроне.
- 5′-граница сплайсинга, расположенная на 5′(левом)-конце интрона, содержит консенсусную последовательность GU.
- 3′-граница сплайсинга, расположенная на 3′(правом)-конце интрона, содержит консенсусную последовательность AG.
- Согласно правилу GU-AG (изначально в контексте ДНК, называемому правилом GT-AG), эти динуклеотиды должны располагаться в двух первых и двух последних позициях всех интронов в последовательности пре-мРНК.

Для точного выяснения молекулярных механизмов сплайсинга необходимо исследовать природу сайтов сплайсинга (границ сплайсинга), последовательностей, непосредственно окружающих сайты разрезания и сшивания.

Сравнивая нуклеотидные последовательности мРНК и соответствующего структурного гена, можно определить местоположение границ экзон–интрон. Анализ нуклеотидного состава показал, что левый и правый концы интрона не имеют протяженных участков гомологии или комплементарности. Место соединения с экзоном, тем не менее, характеризуется наличием коротких консервативных последовательностей.

Специфический конец каждого интрона можно определить, основываясь на консервативности места соединения экзон-интрон. Полученная на основе сравнивания этих участков консенсусная последовательность приведена на РИС. 26.3.

Нижний индекс показывает процент встречаемости определенного основания в каждой конкретной позиции консенсусной последовательности. Высокая консервативность свойственна только тем основаниям, которые, *располагаясь в интроне, непосредственно примыкают* к предполагаемому месту соединения. Таким образом последовательность универсального интрона можно представить в следующем виде:

GU.........AG

Интрон начинается с динуклеотида GU, а заканчивается динуклеотидом AG; места соединения с экзонами в связи с этим описывают на основе **правила GT–AG**. (Подобный способ записи говорит о том, что первоначально анализировали последовательности ДНК, очевидно, что динуклеотид GT в кодирующей цепи ДНК соответствует динуклеотиду GU в последовательности РНК.)

Примечательно, что границы сплайсинга имеют разную последовательность, определяя тем самым *ориентацию* интрона. Их называют слева направо согласно расположению в интроне — 5′-граница (иногда называемая левой или донорной стороной) и 3′-граница (иногда называемая правой или акцепторной стороной). Роль консенсусных последовательностей в распознавании сайтов сплайсинга подтверждается с помощью точечных мутаций, нарушающих сплайсинг *in vivo* и *in vitro*.

26.3 Границы сплайсинга распознаются аппаратом сплайсинга в парах

Ключевые понятия:

- Сплайсинг зависит только от распознавания пар сайтов сплайсинга.
- Все 5′-сайты функционально эквивалентны между собой, равно как неразличимы и все 3′-сайты.

Типичная молекула мРНК млекопитающих содержит в своем составе много интронов. Простота организации сайтов сплайсинга является основной причиной возникновения ошибок. Иллюстрация этой проблемы приведена на РИС. 26.4. Какие процессы лежат в основе правильного сшивания определенных сайтов? Соответствующие GU-AG пары могут быть удалены друг от друга на большие расстояния (длина некоторых интронов превышает 10 тпн). Мы можем предположить два механизма, обеспечивающих корректное спаривание соответствующих 5′- и 3′-сайтов:

- Возможно, способность сшивать сайты, расположенные на концах определенных интронов, является *свойством самой* РНК. Реализация такого механизма



РИС. 26.3 Нуклеотидный состав интронов ядерных генов описывается правилом GU-AG. Нижний индекс показывает процент встречаемости определенного основания в каждой конкретной позиции консенсусной последовательности



РИС. 26.4 Границы сплайсинга определяются только на основе правильных комбинаций консенсусных динуклеотидов на концах интронов

требует узнавания специфических последовательностей или структур.

- Возможно, все 5′-сайты функционально эквивалентны между собой, равно как неразличимы и все 3′-сайты. В этом случае сплайсинг должен протекать по определенным *правилам*, обеспечивающим правильное соединение 5′-сайта со следующим 3′-сайтом.

Из нуклеотидных последовательностей сайтов сплайсинга и окружающих их участков видно, что никакой комплементарности между концами интрона нет; это исключает возможность их непосредственного соединения друг с другом путем спаривания комплементарных оснований. Эксперименты с гибридными предшественниками РНК показали, что любой 5′-сайт может соединиться с любым 3′-сайтом. Так, при соединении первого экзона ранней транскрипционной единицы вируса SV40 с третьим экзоном β-глобинового гена мыши интрон такой гибридной молекулы успешно удаляется при сплайсинге. Такая взаимозаменяемость лежит в основе улавливания экзонов — методики, описанной выше и представленной на рис. 4.11. Все приведенные эксперименты позволяют сделать два основных вывода:

- *Сайты сплайсинга универсальны.* Сайты сплайсинга различных предшественников РНК одинаковы. Индивидуальные предшественники РНК не несут никакой специфичной информации (например, особой вторичной структуры), которая была бы необходима для осуществления сплайсинга.
- *Аппарат сплайсинга не обладает тканевой специфичностью.* Эффективный сплайсинг РНК может происходить в любой клетке, независимо от того, синтезируется ли эта РНК в данной клетке. (Тканеспецифичный альтернативный сплайсинг, который является исключением из этого правила, мы обсудим в разд. 26.12 «Альтернативный сплайсинг — это реализация различных вариантов сшивания экзонов между собой».)

Пародокс заключается в том, что, по-видимому, все 5′-сайты аппарата сплайсинга эквивалентны между собой, в то время как одинаковы и все 3′-сайты. *В принципе любой 5′-сайт может соединиться с любым 3′-сайтом.* Однако в обычных условиях сплайсинг реализуется только между 5′- и 3′-сайтами *одного и того же* интрона. *Какие механизмы обеспечивают распознавание сайтов сплайсинга таким образом, что сплайсинг происходит только между 5′- и 3′-сайтами одного и того же интрона?*

Удаляются ли интроны из РНК в определенном *порядке*? Используя РНК-блоттинг, мы можем выделить ряд промежуточных продуктов, образующихся при сплайсинге ядерных РНК и представляющих собой молекулы, из которых удалены определенные интроны. На РИС. 26.5 приведены результаты анализа овомукоидной мРНК с основными промежуточными продуктами сплайсинга. С помощью блоттинга получен ряд дискретных полос. Это указывает на то, что сплайсинг происходит в определенной последовательности. (Если бы семь интронов удалялись в абсолютно случайном порядке, имелось бы более 300 предшественников с различными комбинациями интронов и никаких дискретных полос не наблюдалось бы.)

Какого-либо *обязательного* порядка сплайсинга, по-видимому, нет, поскольку при анализе могут быть обнаружены промежуточные продуты, у которых интроны удалены в разных комбинациях. Однако имеются

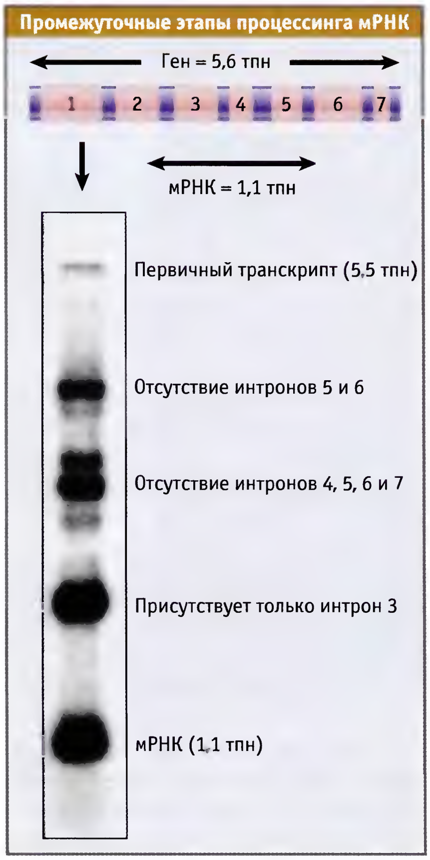


РИС. 26.5 При анализе ядерной РНК с помощью блоттинга по Нозерну и гибридизации с зондом, содержащим овомукоидные последовательности, обнаруживается ряд дискретных предшественников мРНК. Подписи отражают состав основных полосок. Фотография любезно предоставлена Bert W. O'Malley, Baylor College of Medicine

основания предполагать наличие *предпочтительного* порядка протекания сплайсинга. В случае отсутствия одного интрона, им практически всегда является интрон 5 или 6. Каждый из них может быть удален первым. При отсутствии двух интронов, ими также служат чаще всего интроны 5 и 6, однако возможны и другие комбинации. Интрон 3 на первых трех стадиях сплайсинга никогда не удаляется (или удаляется очень редко). Отсюда следует, что имеется предпочтительный порядок сплайсинга, при котором сначала удаляются интроны 5/6, затем интроны 7/4, 2/1, последним удаляется интрон 3. Очевидно, сплайсинг может происходить и в другой последовательности, поскольку (например) в некоторых молекулах последними удаляются интроны 4 или 7. Интерпретировать такие результаты нужно с осторожностью, так как нет данных, доказывающих, что все эти промежуточные продукты превращаются в зрелую мРНК.

Общий вывод, который можно сделать на основе такого анализа, состоит в том, что конформация РНК может влиять на доступность сайтов сплайсинга. При удалении определенных интронов конформация молекулы меняется и становятся доступными новые пары сайтов сплайсинга. Способность предшественника утрачивать интроны более чем в одном порядке позволяет предположить, что молекула на каждой стадии может принимать альтернативные конформации. Очевидно, что с увеличением длины молекулы число таких вариантов увеличивается. Поэтому при рассмотрении генов большего размера становится труднее понять, как определенные вторичные структуры могут контролировать ход реакции. Важный вывод, который следует сделать из такого анализа, заключается в том, что *реакция не распространяется последовательно вдоль предшественника.*

Простой моделью, описывающей механизм контроля распознавания сайтов сплайсинга, могла бы стать модель поступательного движения аппарата сплайсинга. Распознав 5′-сайт, аппарат продолжает движение в определенном направлении вдоль молекулы РНК до тех пор, пока не встретит следующий 3′-сайт. Таким образом может обеспечиваться сплайсинг только между последовательно расположенными сайтами. Эта модель опровергается экспериментами, в которых демонстрируется возможность *транс*-сплайсинга — межмолекулярной реакции, протекающей при создании специальных условий (см. разд. 26.13 «В реакциях *транс*-сплайсинга участвуют малые РНК») либо сплайсинга, происходящего в молекулах РНК, в которых часть нуклеотидной цепи заменена на химический линкер. И в том, и в другом случае нет условий для точного сканирования последовательности РНК по направлению от 5′-сайта сплайсинга к 3′-сайту. Кроме того, данная модель не может объяснить существование явления альтернативного сплайсинга, при котором общий 5′-сайт сшивается с разными 3′-сайтами. Механизмы корректного распознавания сайтов сплайсинга в полной мере пока не выяснены.

26.4 В ходе сплайсинга мРНК образуется структура типа лассо

Ключевые понятия:

- В сплайсинге участвуют три сайта: 5′- и 3′-сайты сплайсинга, а также сайт разветвления, расположенный перед 3′-сайтом сплайсинга.
- Последовательность сайта разветвления высококонсервативна у дрожжей и менее консервативна у высших эукариот.
- После разрезания в 5′-сайте сплайсинга образовавшийся 5′-конец интрона взаимодействует с 2′-положением азотистого основания А, расположенного в сайте разветвления. При этом формируется структуры типа лассо.
- После разрезания в 3′-сайте спласинга интрон высвобождается в форме лассо, а левый и правый экзоны лигируются.
- Сплайсинг реализуется посредством реакций трансэтерификации, в которых связь переносится из одного положения в другое.

In vitro механизмы сплайсинга охарактеризованы с использованием систем, в которых интроны могут уда-

ляться из предшественников РНК. Ядерные экстракты в состоянии обеспечить сплайсинг очищенных предшественников РНК; это свидетельствует о том, что реакция сплайсинга протекает независимо от реакции транскрипции. Сплайсингу может подвергаться некэпированная и неполиаденилированная РНК. Несмотря на то что сплайсинг как таковой не связан с транскрипцией и модификацией РНК, эти процессы скоординированы между собой таким образом, что могут влиять на эффективность сплайсинга.

Этапы сплайсинга *in vitro* проиллюстрированы на РИС. 26.6. Мы обсуждаем реакцию в терминах индивидуальных типов РНК, однако следует помнить, что *in vivo* отдельные виды РНК, содержащие в своем составе экзоны, не высвобождаются в виде отдельных молекул, а удерживаются вместе аппаратом сплайсинга.

На первой стадии сплайсинга происходит разрезание последовательности в 5′-сайте сплайсинга, разделение левого экзона и части молекулы, содержащей в своем составе интрон и правый экзон. Левый экзон сохраняет линейную форму, а правая молекула интрон-экзон формирует **структуру типа лассо**. При этом 5′-конец интрона связывается посредством 5′–2′ связи с нуклеотидом, расположенным внутри интрона. Основанием, с которым взаимодействует 5′-конец молекулы, является А, а последовательность, в состав которой он входит, называют **сайтом разветвления.**

При разрезании в 3′-сайте сплайсинга высвобождается интрон, имеющий форму лассо, а правый экзон сшивается (сплайсируется) с левым экзоном. Для наглядности реакции разрезания и лигирования на рисунке показаны отдельно, однако в действительности они представляют собой один координированный перенос.

После этого лассо разворачивается, интрон приобретает линейную структуру и быстро деградирует.

Последовательности, необходимые для осуществления сплайсинга, представляют собой короткие консенсусные последовательности в 5′- и 3′-сайтах сплайсинга и в сайте разветвления. Принимая во внимание тот факт, что удаление большей части последовательности интрона не препятствует сплайсингу, можно утверждать, что никаких специфических конформаций интрона (или экзона) для успешного сплайсинга не требуется.

Сайт разветвления играет важную роль в распознавании 3′-сайта сплайсинга. Сайт разветвления дрожжей высококонсервативен и представлен консенсусной последовательностью UACUAAC. Сайт разветвления высших эукариот менее консервативен, однако сохраняет состав пуринов и пиримидинов в каждом положении, а также А в качестве нуклеотида — мишени для связи с 5′-концом интрона (см. рис. 26.6).

Сайт разветвления расположен перед 3′-сайтом сплайсинга на расстоянии 18–40 нуклеотидов. У дрожжей мутации или делеции в этом регионе препятствуют сплайсингу. Высшие эукариоты вследствие меньшей консервативности данного региона обладают способностью использовать похожие последовательности (называемые криптическими сайтами) в качестве сайта разветвления в случае нарушения истинного сайта. В этом случае важную роль играет близость такого сайта к 3′-сайту сплайсинга, поскольку криптические сайты всегда расположены рядом с истинным сайтом. Криптические сайты задействуются только в том случае, если истинный сайт инактивирован. При этом сплайсинг протекает по обычному механизму, и образуется тот же конечный продукт, что и в случае дикого типа. *Роль сайта разветвления заключается в идентификации ближайшего 3′-сайта сплайсинга как мишени для связывания с 5′-сайтом сплайсинга.* Это утверждение можно проиллюстрировать фактом взаимодействия между белковыми комплексами, связывающими эти два сайта.

Связь, в результате которой образуется лассо, формируется между 5′-положением инвариантного G, расположенного на 5′-конце интрона, и 2′-положением инвариантного А, расположенного в сайте разветвления (соответствует третьему остатку А в дрожжевой последовательности UACUAAC).

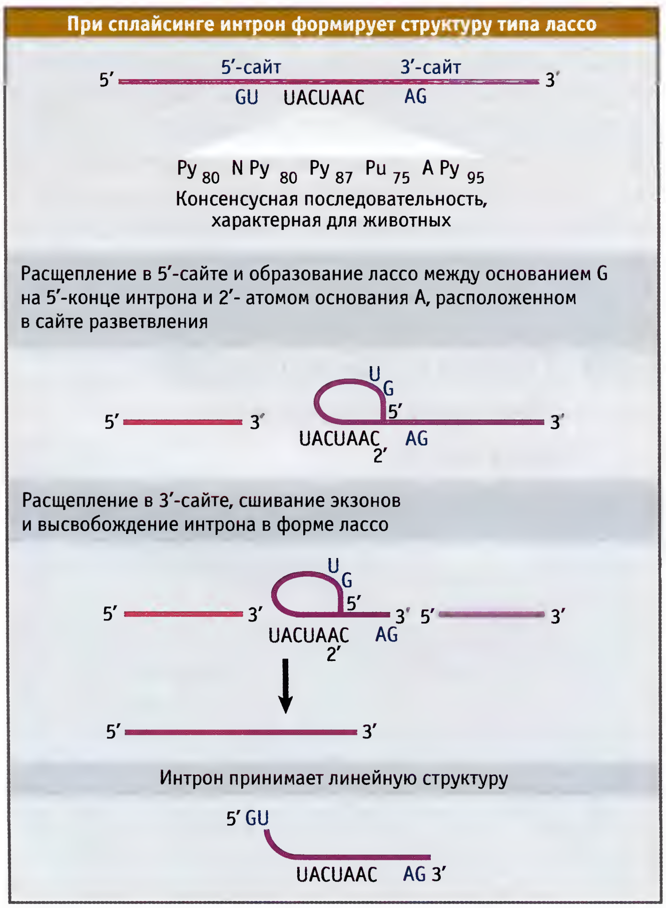


РИС. 26.6 Сплайсинг протекает в две стадии. Сначала происходит отщепление экзона с 5′-конца интрона, а затем присоединение к нему экзона, расположенного с 3′-конца интрона

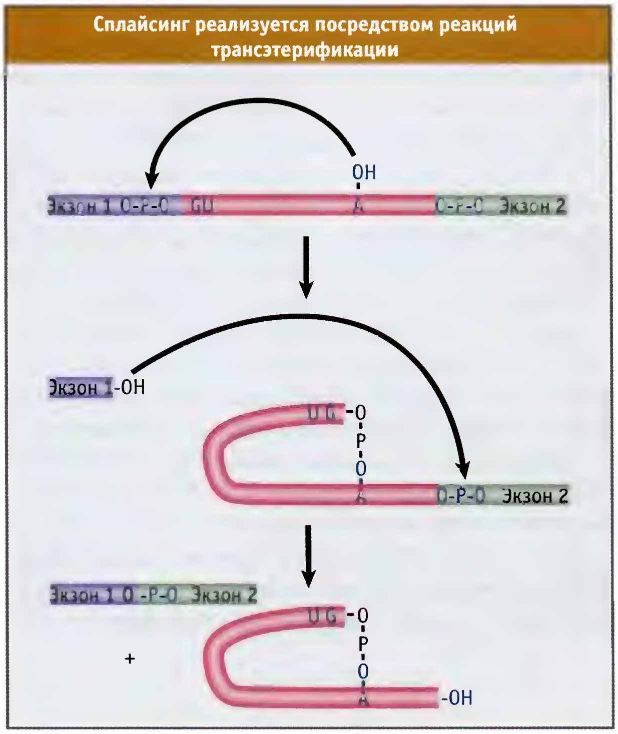


РИС. 26.7 Ядерный сплайсинг реализуется посредством двух реакций трансэтерификации, в ходе которых OH-группы атакуют фосфодиэфирные связи

С химической точки зрения, сплайсинг представляет собой реакцию **трансэтерификации**: связь в действительности *переносится* из одного положения в другое. Как показано на РИС. 26.7, на первом этапе происходит нуклеофильная атака 2′-OH группой инвариантного A (входящего в состав последовательности UACUAAC) 5′-сайта сплайсинга. На втором этапе свободная 3′-OH группа экзона, высвободившегося в результате первой реакции, атакует связь в 3′-сайте сплайсинга. Следует отметить, что число фосфодиэфирных связей постоянно. Первоначальные две 5′–3′ связи на границах сплайсинга экзон-интрон заменяются на одну 5′–3′ связь между экзонами и одну 5′–2′ связь, формирующую лассо.

26.5 мяРНК участвуют в сплайсинге

Ключевые понятия:

- К пяти мяРНК, вовлеченным в процесс сплайсинга, относятся U1, U2, U5, U4 и U6.
- Вместе с несколькими дополнительными белками мяРНК формируют сплайсосому.
- За исключением U6, все мяРНК содержат консервативную последовательность, связывающую Sm-белки, распознаваемые антителами, образующимися при аутоиммунном заболевании.

Компоненты аппарата сплайсинга, участвующие в распознавании 5′- и 3′-сайтов сплайсинга, формируют большой комплекс. Этот комплекс сводит вместе 5′- и 3′-сайты сплайсинга еще до начала реакции. Данный факт объясняет, почему отсутствие любого из сайтов препятствует инициации реакции. Последовательная сборка комплекса происходит на пре-мРНК через ряд промежуточных стадий, о чем свидетельствует обнаружение комплексов разных размеров при фракционировании. Сплайсинг инициируется только после того, как все компоненты будут собраны.

Аппарат сплайсинга представлен как белками, так и РНК (помимо пре-мРНК). Молекулы малых РНК существуют в виде рибонуклеопротеиновых частиц. И ядро, и цитоплазма эукариотических клеток содержат большое разнообразие видов малых РНК. Их размер варьирует от 100 до 300 нуклеотидов у высших эукариот и достигает порядка 1000 оснований у дрожжей. Число молекул малых РНК колеблется от 10^5–10^6 на клетку до количеств, не обнаруживаемых прямыми методами детекции.

Молекулы, локализующиеся в ядре, называют **малыми ядерными РНК (мяРНК)**; аналогичные РНК, обнаруживаемые в цитоплазме, называют **малыми цитоплазматическими РНК (мцРНК)**. Природные малые РНК существуют в виде рибонуклеопротеиновых частиц (мяРНП и мцРНП). В англоязычных лабораториях такие РНК иногда называют **snurps** и **scyrps**. В ядре обнаружен еще один класс малых РНК, названных snoРНК и участвующих в процессинге рибосомальной РНК (см. разд. 26.22 «Малые РНК участвуют в процессинге рРНК»).

мяРНК, задействованные в сплайсинге вместе с дополнительными белками, формируют большой комплекс, названный **сплайсосомой**. Такой комплекс, выделенный из систем сплайсинга *in vitro*, представляет собой рибонуклеопротеиновую частицу с коэффициентом седиментации от 50S до 60S. Процесс формирования сплайсосомы проходит ряд промежуточных стадий, на которых образуется несколько «пресплайсинговых комплексов». Сплайсосома представляет собой большую частицу, превосходящую по массе рибосому.

На РИС. 26.8 схематично изображены все компоненты сплайсосомы. На 5 мяРНК приходится больше четверти всей массы комплекса; вместе с ассоциированными с ними белками (41 белок) на их долю приходится почти половина всей массы частицы. 70 других белков, обнаруженных в сплайсосоме, описаны как факторы сплайсинга. Сюда включены белки, необходимые для сборки сплайсосомы, для связывания РНК-субстрата и белков, вовлеченных в каталитический процесс. Кроме того, еще порядка 30 белков, ассоциированных со сплайсосомой, участвуют в регуляции экспрессии генов на других стадиях. Это позволяет предположить, что сплайсосома может выступать в роли координирующего аппарата.

Сплайсосома формируется на интактном предшественнике РНК и остается на нем во время отщепления



РИС. 26.8 Масса сплайсосомы составляет примерно 12 МДа. Пять мяРНП составляют почти половину массы комплекса. Оставшиеся белки представлены факторами сплайсинга и белками, вовлеченными в другие стадии экспрессии генов

линейной молекулы левого экзона от правой молекулы лассо-интрон-экзон. В комплексе обнаруживается лишь небольшое количество продукта сплайсинга. Это означает, что он, как правило, высвобождается немедленно после расщепления в 3′-сайте и лигирования экзонов.

мяРНП формируют структуру сплайсосомы. Как и в рибосоме, структура комплекса зависит от РНК-РНКовых, РНК-белковых и белок-белковых взаимодействий. В ряде реакций, в которых задействованы мяРНП, происходит прямое спаривание входящих в их состав РНК с последовательностью РНК, подвергающейся сплайсингу; другие реакции обеспечиваются распознаванием между мяРНП или между ассоциированными с ними белками и другими компонентами сплайсосомы.

Значимость молекул мяРНК может быть продемонстрирована с помощью мутаций, вводимых в их гены у дрожжей. Мутации в генах 5 мяРНК летальны и блокируют сплайсинг. Все мяРНК, задействованные в сплайсинге, консервативны и обнаруживаются в неизменном виде в клетках животных, птиц и насекомых. Соответствующие РНК дрожжей часто существенно больше, однако консервативные регионы обладают теми же характеристиками, что и мяРНК высших эукариот.

U1, U2, U5, U4 и U6 — мяРНП, участвующие в сплайсинге. Свое название они получили в соответствии с мяРНК, входящей в их состав. Каждая мяРНП содержит единственную мяРНК и несколько (<20) белков. U4 и U6 мяРНП чаще всего обнаруживаются в составе одной частицы (U4/U6). Общее структурное ядро всех мяРНП представлено восемью белками, распознаваемыми аутоиммунной анти-сывороткой — **анти-Sm**. Консервативные участки этих белков служат мишенью для взаимодействия с антителами. Остальные белки, входящие в состав каждой мяРНП, уникальны для нее. Sm белки связывают консервативную последовательность $PuAU_{3-6}GPu$, присутствующую во всех мяРНП, за исключением U6. U6, в свою очередь, содержит в своем составе группу Sm-подобных белков. Sm белки вовлечены в патогенез аутоиммунных реакций, хотя их точная роль в этих процессах не выяснена.

Некоторые белки мяРНП могут участвовать в сплайсинге напрямую; другие играют лишь структурную роль и необходимы для сборки или взаимодействия между мяРНП. Около трети всех белков, задействованных в сплайсинге, являются компонентами мяРНП. Свидетельства прямой роли РНК в сплайсинге позволяют предположить, что только несколько белковых факторов сплайсинга напрямую задействованы в катализе; основная их часть обеспечивает сборку частиц и выполняет структурную функцию.

26.6 U1 мяРНП инициирует сплайсинг

Ключевые понятия:

- U1 мяРНП инициирует сплайсинг, связываясь с 5′-сайтом сплайсинга с помощью РНК-РНК взаимодействий.
- В состав комплекса E входит U1 мяРНП, связанная с 5′-сайтом сплайсинга, белок U2AF, связанный с пиримидиновым регионом между сайтом разветвления и 3′-сайтом сплайсинга, и группа SR белков, соединяющих U1 мяРНП с U2AF.

Реакция сплайсинга может быть разделена на две стадии:

- Сначала распознаются консенсусные последовательности 5′-сайта сплайсинга, сайта разветвления и примыкающего пиримидинового региона. Происходит сборка комплекса, включающего в себя все компоненты сплайсинга.
- Реакции разрезания и лигирования изменяют структуру субстратной РНК. По окончанию реакции сплайсинга компоненты комплекса высвобождаются или реорганизуются.

До тех пор пока все компоненты сплайсинга не будут собраны, а все сайты сплайсинга не станут доступными, субстратная РНК не подвергается никаким необратимым изменениям.

В распознавании консесусных последовательностей задействованы как РНК, так и белки. Отдельные мяРНК содержат в своем составе последовательности, комплементарные консенсусным последовательностям пре-мРНК или других мяРНК. Спаривание между мяРНК и пре-мРНК или мяРНК между собой играет в сплайсинге важную роль.

U1 мяРНП человека содержит в своем составе восемь белков и РНК-компонент. Вторичная структура

U1 мяРНП представлена на РИС. 26.9. Она образует несколько доменов. Sm-связывающий сайт обеспечивает взаимодействие с общими белками мяРНП. Домены, формирующие структуры типа стебель-петля, служат сайтами связывания с белками, уникальными для U1 мяРНП.

Связывание U1 мяРНП с 5′-сайтом сплайсинга является первой стадией реакции сплайсинга. При этом один из белков U1 мяРНП (U1-70k) взаимодействует с белком ASF/SF2 (основным фактором сплайсинга SR-класса). U1 мяРНП спаривается с 5′-сайтом сплайсинга с помощью одноцепочечной 5′-концевой последовательности, которая обычно представлена четырьмя–шестью основаниями, комплементарными сайту сплайсинга на пре-мРНК.

Роль взаимодействия между 5′-сайтом сплайсинга и U1 мяРНП может быть оценена с помощью мутаций в этих регионах. Результат такого эксперимента приведен на РИС. 26.10. Сайт сплайсинга 12S аденовирусной пре-мРНК с последовательностью дикого типа спаривается пятью (из шести) основаниями с U1 мяРНП. Мутантная 12S РНК с двумя нуклеотидными заменами не может быть сплайсирована; остатки GG в положениях интрона 5 и 6 заменены на остатки AU. В результате мутаций изменяется характер спаривания между U1 мяРНП и 5′-сайтом сплайсинга, хотя *общая* протяженность спаренного региона остается прежней (потеря комплементарности в одной позиции и приобретение ее в другой). Эффект, оказываемый этими изменениями на сплайсинг, свидетельствует о важности таких взаимодействий.

При введении в U1 мяРНП мутации, восстанавливающей спаривание в позиции 5, сплайсинг возобновляется. На основе явления супрессии мутаций в 5′-сайте сплайсинга мутациями в U1 мяРНП можно вывести общее правило: комплементарность между U1 мяРНП и 5′-сайтом сплайсинга является необходимым условием сплайсинга, однако эффективность сплайсинга определяется не только числом образующихся пар оснований. Реакция спаривания стабилизируется белками U1 и мяРНП.

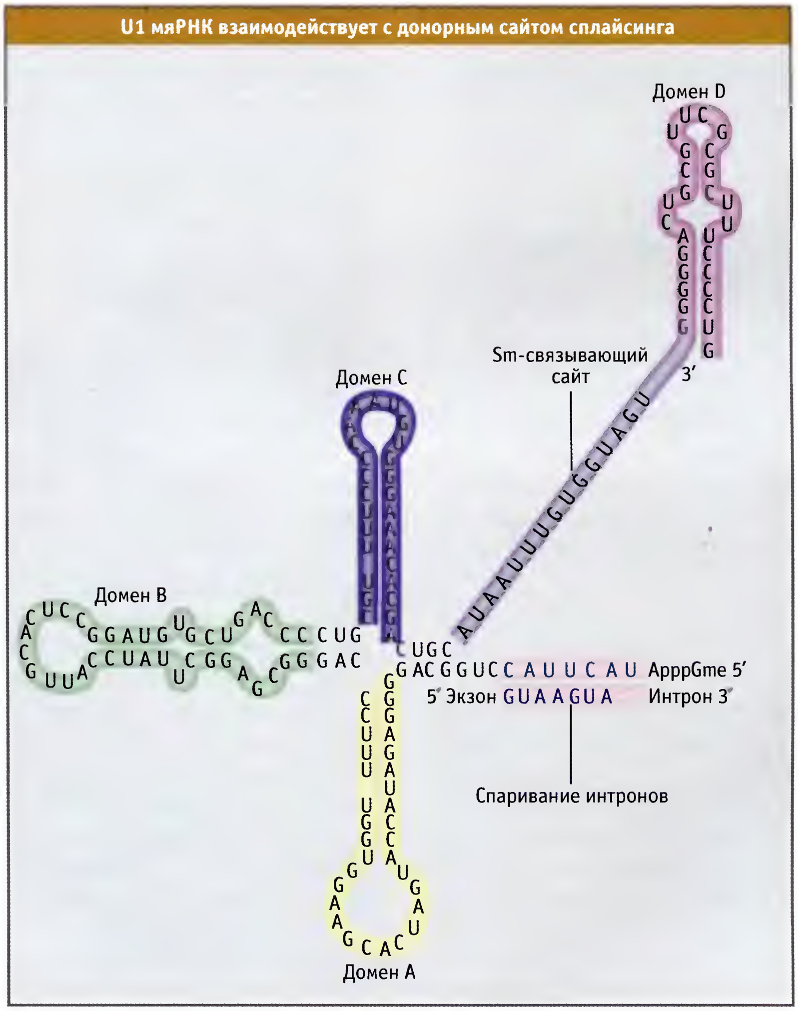


РИС. 26.9 Вторичная структура U1 мяРНК представлена несколькими доменами. Ее 5′-конец остается одноцепочечным и может взаимодействовать с 5′-сайтом сплайсинга

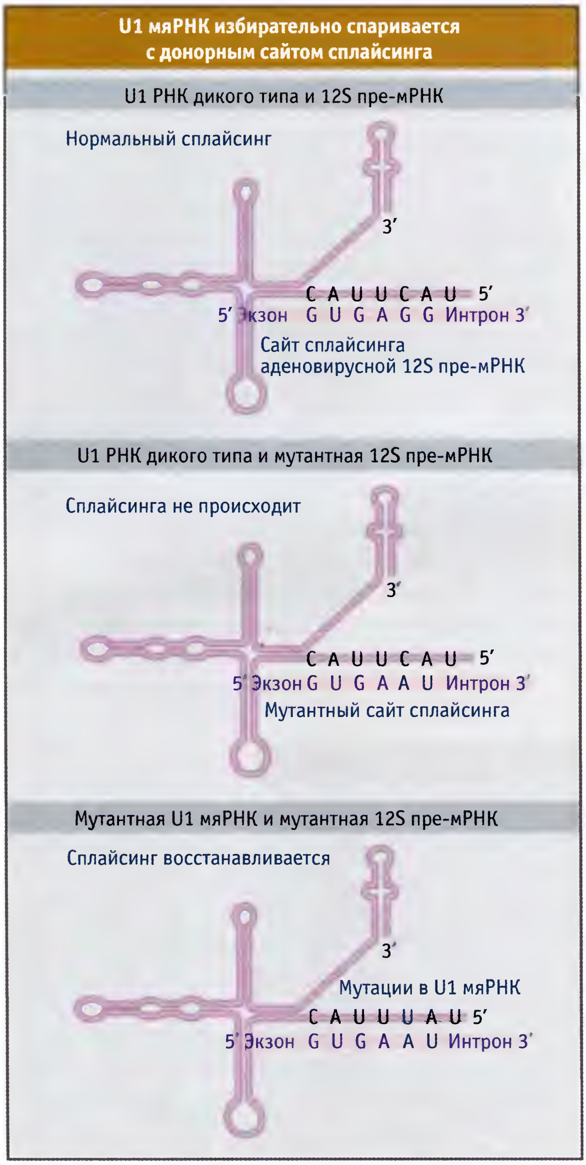


РИС. 26.10 Мутации, нарушающие функцию 5′-сайта сплайсинга, могут быть супрессированы в результате компенсаторных мутаций в U1 мяРНК, восстанавливающих корректное спаривание оснований

На РИС. 26.11 изображены начальные стадии сплайсинга. Первым формируется комплекс E (ранний пресплайсинговый комплекс), объединяющий U1 мяРНП, фактор сплайсинга U2AF, и белки — члены **семейства SR**, представляющие собой важную группу факторов сплайсинга и регуляторных факторов. Свое название они получили благодаря наличию регионов, богатых аргинином и серином. Эти белки обладают способностью связывать РНК. Являясь необходимым компонентом сплайсосомы, они формируют каркас комплекса на РНК-субстрате, связывая U2AF и U1 (РИС. 26.12). Комплекс E иногда называют фиксационным комплексом, поскольку его образование определяет молекулу пре-мРНК как субстрат для формирования комплекса сплайсинга.

U2AF, входящая в состав комплекса E, связана с регионом пре-мРНК, расположенным между сайтом разветвления и 3′-сайтом сплайсинга. Свое название белок U2AF получил в результате того, что первоначально был определен как вспомогательный фактор мяРНП U2. У большинства организмов он состоит из большой субъединицы (U2AF65), взаимодействующей с пиримидиновым участком, расположенным после сайта разветвления, и малой субъединицы (U2AF35), которая связывается с динуклеотидом AG в 3′-сайте сплайсинга. У *Saccharomyces cerevisiae* эта функция выполняется белком Mud2, который является аналогом U2AF65 и связывается только с пиримидиновым регионом. Этой особенностью обусловливается различие механизмов сплайсинга *S. cerevisiae* и других организмов. У дрожжей 3′-сайт сплайсинга не участвует в формировании комплекса сплайсинга на ранних стадиях реакции, в то время как у всех остальных организмов 3′-сайт сплайсинга в этот процесс вовлечен.

Другой фактор сплайсинга, названный SF1 у млекопитающих и BBP у дрожжей, присоединяет U2AF/Mud2 к U1 мяРНП, связанной с 5′-сайтом сплайсинга. Структура комплекса упрочняется за счет совместного действия двух белков; SF1 и U2AF (или BBP и Mud2) связывают РНК-субстрат вместе приблизительно в 10 раз эффективнее, чем по отдельности. Возможно, это взаимодействие обеспечивает первый контакт двух сайтов сплайсинга.

Комплекс E превращается в комплекс A в тот момент, когда мяРНП U2 связывается с сайтом разветвления. И U1 мяРНП, и U2AF/Mud2 должны связаться с U2. Эта мяРНП содержит последовательность, комплементарную сайту разветвления. Последовательность, расположенная рядом с 5′-концом мяРНП, спаривается с сайтом разветвления интрона. У дрожжей при этом, как правило, образуется дуплекс, содержащий в своем составе участок UACUAAC (см. рис. 26.14). Несколько белков U2 мяРНП связываются с РНК-субстратом непосредственно перед сайтом разветвления. При присоединении U2 мяРНП к комплексу E формируется пресплайсинговый комплекс A. Связывание U2 мяРНП — АТФ-зависимый процесс, обеспечивающий старт реакции сплайсинга.

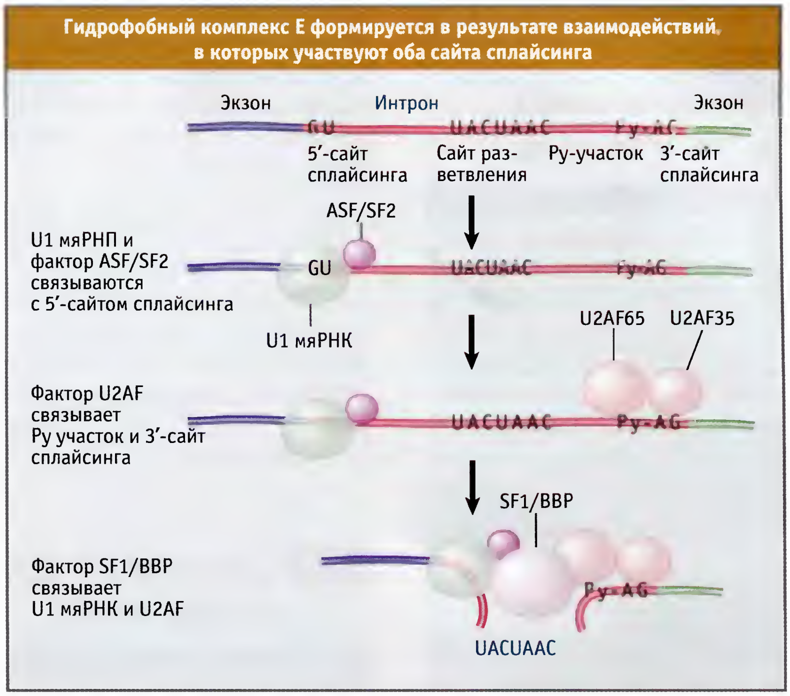


РИС. 26.11 Комплекс E формируется в результате взаимодействия U1 мяРНП с 5′-сайтом сплайсинга, а фактора U2AF — с пиримидиновым участком и 3′-сайтом сплайсинга с последующим присоединением U1 мяРНП к U2AF с помощью белка SF1/BBP

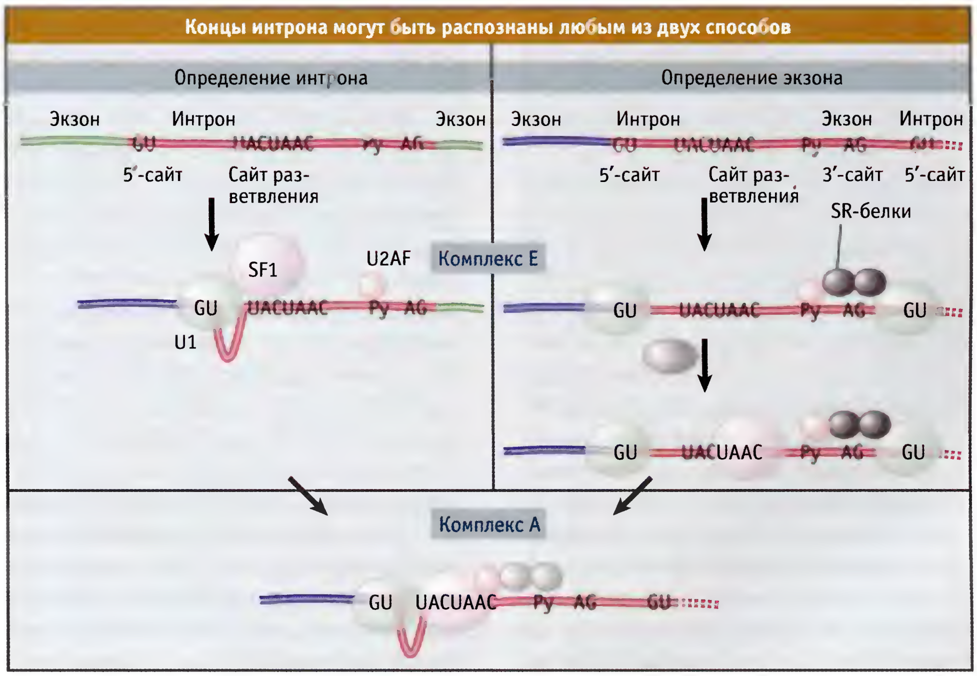


РИС. 26.12 Первичное распознавание 5′- и 3′-сайтов сплайсинга может происходить разными путями

26.7 Формирование комплекса E может происходить в результате распознавания как интронов, так и экзонов

Ключевые понятия:

- Прямой путь формирования комплекса E представляет собой связывание U1 мяРНП с 5′-сайтом сплайсинга, и U2AF — с пиримидиновым регионом, расположенным между сайтом разветвления и 3′-сайтом сплайсинга.
- Другая возможность заключается во взаимодействии U2AF с пиримидиновым регионом и U1 мяРНП с 5′-сайтом сплайсинга следующего интрона.
- Комплекс E превращается в комплекс A в результате связывания U2 мяРНП с сайтом разветвления.
- В случае когда комплекс E формируется при участии 5′-сайта сплайсинга следующего интрона, при превращении комплекса E в комплекс A этот сайт сплайсинга заменяется правильным 5′-сайтом.
- Слабый 3′-сайт сплайсинга требует участия энхансеров сплайсинга — последовательностей, расположенных в правом экзоне и способных напрямую связываться с SR-белками.

Формирование комплекса E может происходить разными путями. Возможные варианты этого процесса приведены на рис. 26.12. Наиболее простым является распознавание сайтов сплайсинга, расположенных в одном интроне. Присутствие U2 мяРНП в 5′-сайте сплайсинга является условием для связывания U2AF с пиримидиновым регионом, расположенным после сайта разветвления. При этом обеспечивается контакт 5′- и 3′-концов интрона между собой. Комплекс E превращается в комплекс A в результате связывания U2 мяРНП с сайтом разветвления. *Основная особенность этого пути сплайсинга заключается в том, что распознавание обоих сайтов сплайсинга не требует участия последовательностей, расположенных за пределами интрона*. Этот процесс был назван **распознаванием интрона**.

В исключительных случаях, когда SR-белки становятся неспособны связать U2AF/U2 *in vitro* (в отсутствие U1 мяРНП), увеличивается возможность включения механизма U1-независимого сплайсинга.

Альтернативный путь формирования сплайсосомы может быть задействован в том случае, если интроны слишком длинные, а сайты сплайсинга слабые. Как показано в правой части рисунка, 5′-сайт сплайсинга распознается U1 мяРНП обычным образом. 3′-сайт сплайсинга распознается как часть комплекса, формирующегося в *следующем экзоне*, в котором 5′-сайт сплайсинга тоже связан с U1 мяРНП. U1 мяРНП соединяется с U2AF в пиримидиновом регионе с помощью SR-белков. При присоединении U2 мяРНП и формировании комплекса A, в результате перестроек корректный (в большинстве случаев левый) 5′-сайт замещается в комплексе следующим 5′-сайтом. Важная особенность этого механизма сплайсинга заключается в том, что в нем задействуются последовательности, расположенные за пределами интрона. Обычно эти последовательности включают в себя следующий 5′-сайт сплайсинга. Этот процесс был назван **распознаванием экзона.** Данный механизм не универсален: ни белки SR, ни распознавание экзона не характерно для *S. cerevisiae.*

«Слабые» 3′-сайты не способны эффективно связывать U2AF и U2 мяРНП. Для взаимодействия с SR-белками необходимы дополнительные последовательности. Такие последовательности называют «энхансерами сплайсинга». Они обеспечивают связывание U2AF в пиримидиновом регионе. Как правило, эти последовательности расположены в экзоне после 3′-сайта сплайсинга.

26.8 Пять мяРНП формируют сплайсосому

Ключевые понятия:

- При связывании U5 и U4/U6 мяРНП комплекс A превращается в B1 сплайсосому, которая уже содержит все компоненты, необходимые для сплайсинга.
- Во время сплайсинга сплайсосома подвергается изменениям, последовательно проходя несколько стадий.
- Высвобождение U1 мяРНП дает возможность U6 мяРНП взаимодействовать с 5′-сайтом сплайсинга и обеспечивает превращение сплайсосомы B1 в сплайсосому B2.
- Диссоциация U4 мяРНП от U6 мяРНП приводит к спариванию U6 с U2 и формированию каталитически активного сайта.

Вслед за формированием комплекса E все остальные мяРНП и факторы, участвующие в сплайсинге, вовлекаются в процесс в определенном порядке. На рис. 26.13 представлены все компоненты, которые можно последовательно идентифицировать в ходе реакции.

Комплекс B1 формируется, когда тример, состоящий из U5 и U4/U6 мяРНП, связывается с комплексом A, состоящим из U1 и U2 мяРНП. Этот комплекс уже рассматривается как сплайсосома, поскольку содержит в своем составе все компоненты, необходимые для реакции сплайсинга. После высвобождения U1 мяРНП сплайсосома B1 превращается в комплекс B2. Диссоциация U1 мяРНП делает возможным непосредственное сближение остальных компонентов реакции (в особенности U6 мяРНП) с 5′-сайтом сплайсинга. Изменяет свое положение и U5 мяРНП: первоначально U5 располагается рядом с последовательностью экзона в области 5′-сайта сплайсинга, однако затем сдвигается к интрону.

Каталитическая реакция запускается при высвобождении U4; это — АТФ-зависимый процесс. Роль U4 мяРНП заключается в блокировании U6 мяРНП до того момента, пока она не понадобится. На рис. 26.14 показано, как изменяется спаривание между элементами мяРНП во время сплайсинга. Протяженный участок U6 мяРНП длиной 26 оснований спаривается с двумя отдельными регионами U4 мяРНП. После диссоциации

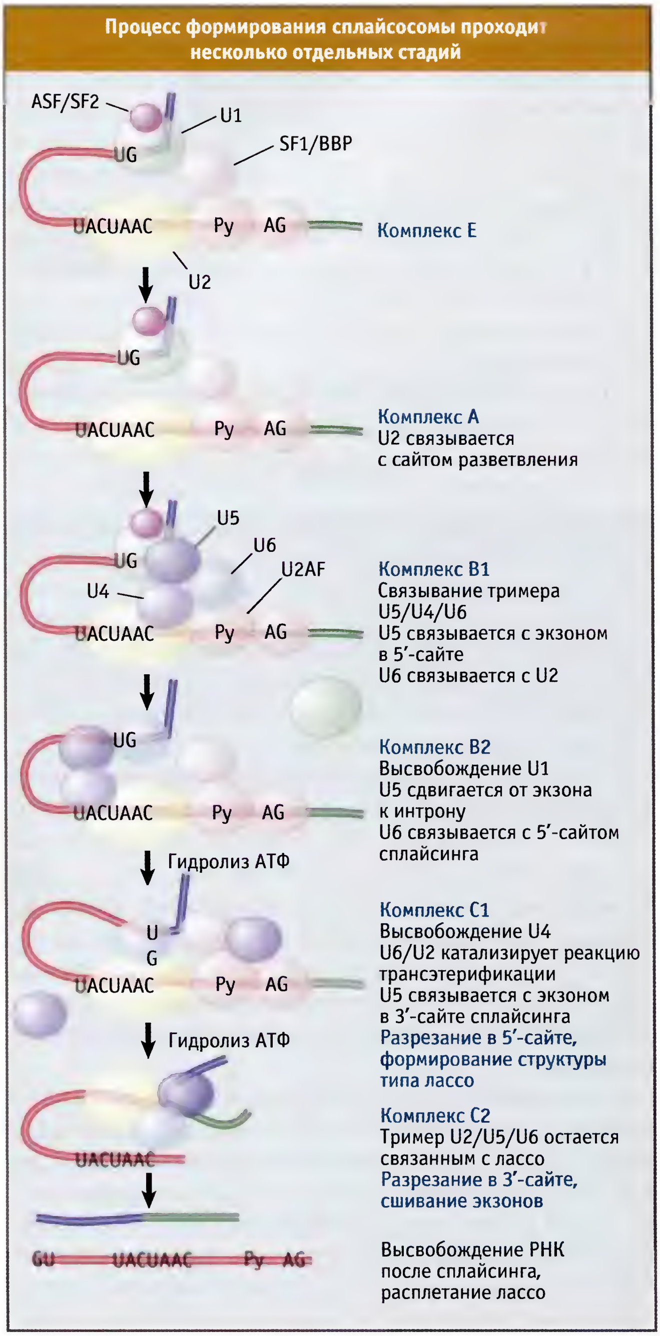


РИС. 26.13 Реакция сплайсинга протекает в несколько стадий. Формирование сплайсосомы происходит при взаимодействии компонентов, распознающих консенсусные последовательности

U4 освободившаяся последовательность U6 принимает другую структуру. Первая ее часть спаривается с U2; вторая формирует внутримолекулярную шпильку. Взаимодействие между U4 и U6 исключает взаимодействие между U2 и U6, поэтому высвобождение U4 играет регуляторную роль, контролируя способность сплайсосомы осуществлять реакцию далее.

Для наглядности РНК-субстрат на рисунке изображен в протяженной форме, однако в действительности 5′-сайт сплайсинга расположен в непосредственной близости от последовательности U6, находящейся с 5′-стороны участка молекулы, связанного с U2. Последовательность U6 мяРНП спаривается с последовательностью интрона за консервативным динуклеотидом GU в 5′-сайте сплайсинга (мутации, повышающие прочность этого спаривания, увеличивают эффективность сплайсинга).

Во время сплайсинга происходит несколько реакций спаривания между мяРНП и субстратной РНК. Все они изображены на РИС. 26.15. мяРНП содержат в своем составе последовательности, комплементарные последовательностям субстрата и друг другу. Кроме того они имеют одноцепочечные участки, формирующие петли. Располагаясь в непосредственной близости от субстрата, эти последовательности играют в сплайсинге важную роль, что доказывается способностью мутаций в этих регионах блокировать сплайсинг.

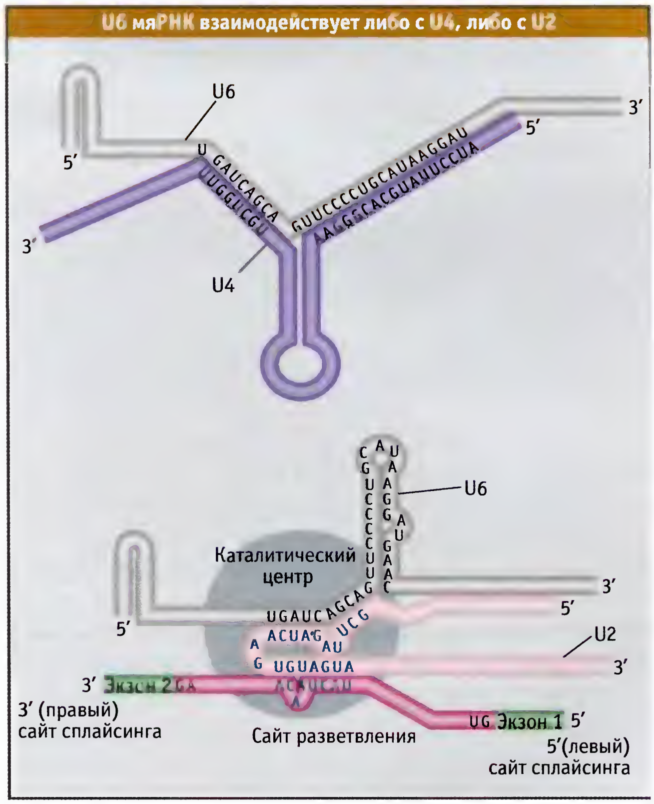


РИС. 26.14 Взаимодействия U6-U4 и U6-U2 являются взаимоисключающими. В момент присоединения к сплайсосоме U6 спарен с U4. Высвобождение U4 вызывает конформационные изменения в структуре U6; одна часть освободившейся последовательности формирует шпильку (обозначена темно-серым цветом), другая (обозначена черным цветом) – спаривается с U2. Регион U2, взаимодействующий с U6, к тому моменту уже оказывается спаренным с сайтом разветвления, вследствие чего U6 оказывается в непосредственной близости от точки рзветвления. Следует обратить внимание на то, что субстратная РНК на рисунке изображена в необычном положении и ориентирована в направлении от 3′- к 5′-концу

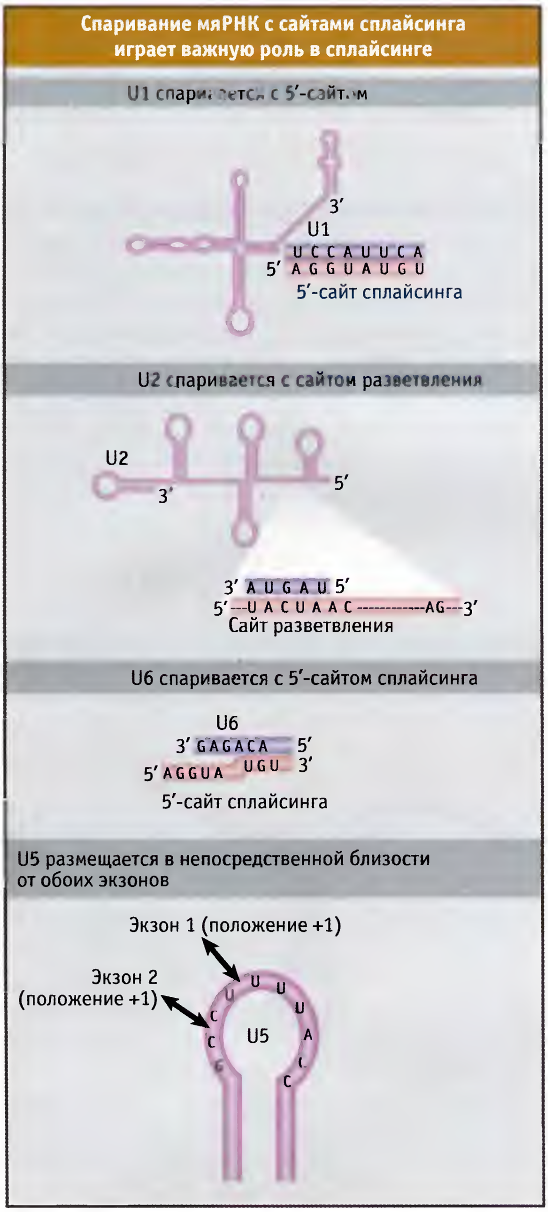


РИС. 26.15 **Во время сплайсинга между сайтами сплайсинга и мяРНК происходит серия реакций спаривания оснований**

В результате спаривания между U2 и последовательностью сайта разветвления, а также между U2 и U6 формируется структура, напоминающая активный центр способных к автосплайсингу интронов группы II (см. рис. 26.20). Исходя из этого, можно предположить, что основу каталитической компоненты составляет РНК-структура, образованная в результате взаимодействия между U2 и U6. U6 спаривается с 5'-сайтом сплайсинга, а эксперименты с перекрестным связыванием показали, что петля U5 мяРНК примыкает к основаниям, расположенным в первом положении каждого экзона. Мы можем оценить местоположение субстрата (5'-сайта и сайта связывания), мяРНК (U2 и U6) относительно каталитического центра (как показано на рис. 26.14), однако компоненты, ответственные за траснэтерификацию, не были определены напрямую.

Формирование лассо в сайте разветвления обеспечивает включение в реакцию 3'-сайта сплайсинга, поскольку 3'-консенсусная последовательность, расположенная с 3'-стороны сайта разветвления становится мишенью для второй реакции трансэтерификации. Вторая стадия сплайсинга протекает быстро. Несмотря на то что для ее осуществления необходимо связывание U5 мяРНП с 3'-сайтом сплайсинга, свидетельств того, что это взаимодействие опосредовано комплементарным спариванием, нет.

Важное замечание, которое можно сделать на основе этих результатов, заключается в том, что *мяРНК- компоненты аппарата сплайсинга взаимодействуют друг с другом и с субстратной РНК посредством комплементарного спаривания. Эти взаимодействия приводят к структурным изменениям, в результате которых реакционные группы сближаются, формируя каталитические центры.* Более того, конформационные изменения в структуре мяРНК обратимы; к примеру, U6 мяРНК, не использованная в реакции сплайсинга, освободившись от U2, способна восстановить дуплекс с U4 и вступить в новый цикл сплайсинга.

Мы описали реакции, в которых участвуют отдельные мяРНП, тем не менее, возможно, каждая мяРНП играет в сплайсинге несколько ролей, участвуя в серии реакций. Так, например, U1 мяРНП участвует в двух независимых процессах: способствует связыванию U2 мяРНП с сайтом разветвления и сама связывается с 5'-сайтом сплайсинга. Похожим образом разные регионы U2 мяРНК задействованы в связывании с сайтом разветвления и во взаимодействии с другими компонентами сплайсинга.

Обширный мутационный анализ компонентов сплайсинга у дрожжей позволил выявить рибонуклеиновые и белковые составляющие сплайсосомы. Мутации в генах, необходимых для сплайсинга, определялись по накоплению несплайсированных предшественников мРНК. Ряд локусов, содержащих гены, кодирующие задействованные в сплайсинге белки, и первоначально называемые *РНК,* теперь известны как *PRP* (пре-РНК-процессинговые). Некоторые продукты этих генов содержат мотивы, наличие которых позволяет отнести их группе РНК-связывающих белков, а некоторые являются АТФ-зависимыми РНК-хеликазами. Мы предполагаем, что, помимо РНК-РНК взаимодействий, РНК-белковые взаимодействия тоже играют важную роль в создании или расформировании структур пре-мРНКовых и мяРНКовых компонентов сплайсосом.

Некоторые из белков PRP являются компонентами мяРНП, другие функционируют как независимые факторы. Интересным примером является PRP16 — хеликаза, гидролизующая АТФ и кратковременно связывающаяся со сплайсосомой на второй стадии каталитической реакции. Другим примером является PRP22 — еще одна АТФ-зависимая хеликаза, необходимая для высвобождения зрелой мРНК из сплайсосомы. Сохранение коли-

чества связей в ходе сплайсинга означает, что для формирования связей привнесения энергии не требуется, а это, в свою очередь, значит, что энергия гидролиза АТФ расходуется на другие цели. Использование АТФ PRP16 и PRP22 может служить примером общего феномена использования гидролиза АТФ для запуска конформационных изменений, необходимых для осуществления сплайсинга.

26.9 Аппарат альтернативного сплайсинга использует различные мяРНП

Ключевые понятия:

- В реакции альтернативного сплайсинга задействован другой набор мяРНП. Комплекс, который они формируют, называют сплайсосомой U12.
- Интроны-мишени можно идентифицировать с помощью длинных консенсусных последовательностей на границах сплайсинга, однако обычно они содержат те же самые GU-AG динуклеотиды.
- В некоторых интронах на границах сплайсинга расположены динуклеотиды AU-AC. Часть из этих интронов U1-зависимые, часть — U12-зависимые.

Интроны GU-AG составляют подавляющее большинство интронов (>98% границ сплайсинга в человеческом геноме). На границах лишь менее 1% интронов расположены GC-AG динуклеотиды. Кроме того, существует еще один небольшой класс интронов (0,1% всех интронов), начинающихся и оканчивающихся динуклеотидами AU-AC. Первые из этих интронов были открыты как компоненты аппарата альтернативного сплайсинга — сплайсосомы U12. Сплайсосома U12 состоит из U11 и U12 мяРНП (соответствуют U1 и U2 мяРНП), варианта мяРНП U5, а также $U4_{atac}$ и $U6_{atac}$ мяРНП. Сама реакция сплайсинга протекает по тем же механизмам, что и сплайсинг GU-AG интронов, и мяРНП играют аналогичные роли. Существует ли разница между белковыми компонентами этих аппаратов, неизвестно.

К настоящему моменту выяснено, что последовательность интрона влияет на тип формирующейся сплайсосомы. Некоторые AU-AC интроны сплайсируются сплайсосомами U2, некоторые GU-AG интроны сплайсируются сплайсосомами U12. Строгая консенсусная последовательность с левого края интрона направляет сплайсинг по U12-зависимому типу:

$$5'{}^{G}{}_{A}UAUCCUUU\ldots\ldots PyA^{G}{}_{C}3'.$$

По концам большинства U12-зависимых интронов находятся динукледотиды GU……AG. Кроме того, они характеризуются высококонсервативным сайтом разветвления — UCCUUPuAPy, который спаривается с U12. По этим причинам, термин U12-зависимый интрон является синонимом термина AU-AC интрон.

Два типа интронов сосуществуют во многих геномах, а в некоторых случаях даже обнаруживаются в одном гене. U12-зависимые интроны, как правило, фланкированы U2-зависимыми интронами. Данные о филогении этих интронов позволяют предположить, что когда-то AU-AC U12-зависимые интроны были шире распространены, но в ходе эволюции концевые динуклеотиды AU-AC сменились на GU-AG, и сплайсинг приобрел U2-зависимый характер. На общую эволюцию этих систем указывает и аналогичный характер спаривания мяРНП с субстратной пре-мРНК, и мяРНП между собой.

Задействование мяРНП в реакциях сплайсинга является лишь одним из примеров их участия в процессинге РНК. мяРНП необходимы при превращении ядерной РНК в зрелую рРНК. В свете таких фактов, как способность интронов I группы к автосплайсингу и наличие каталитической активности у РНК компонента рибонуклеазы Р (рассматривается в гл. 27 «Каталитическая РНК»), можно утверждать, что РНК-РНК взаимодействия играют большую роль во многих событиях процессинга РНК.

26.10 Процесс сплайсинга сопряжен с экспортом мРНК

Ключевые понятия:

- Белки REF связывают сайты сплайсинга, взаимодействуя со сплайсосомой.
- После окончания сплайсинга они остаются связанными с РНК в области границы экзон-экзон.
- Белки REF взаимодействуют с транспортным белком TAP/Mex, экспортирующим РНК сквозь ядерную пору.

После окончания синтеза и процессинга мРНК экспортируется из ядра в цитоплазму в составе рибонуклеопротеинового комплекса. Белки, ответственные за транспорт между ядром и цитоплазмой, остаются в компартменте очень непродолжительное время. Интерес представляют два момента: как эти белки распознают свой РНК-субстрат и какие механизмы обеспечивают транспорт только полностью процессированной мРНК. Возможно, объяснение кроется в четком координировании всех событий: сплайсосома может формироваться до того, как заканчивается транскрипция. Может существовать и прямая связь между сплайсингом и экспортом мРНК.

Интроны могут препятствовать экспорту мРНК, будучи связанными с аппаратом сплайсинга. Сплайсосома же может обеспечивать первичный контакт мРНК и транспортного аппарата. На рис. 26.1[illegible] показана модель, в которой белковый комплекс связывается с РНК при помощи аппарата сплайсинга. Комплекс состоит более чем из 9 белков и называется комплексом соединения экзонов (КСЭ).

Комплекс КСЭ участвует в реализации некоторых функций сплайсированных мРНК. Некоторые компоненты комплекса напрямую вовлечены в выполнение этих функций, другие — требуют участия дополнительных

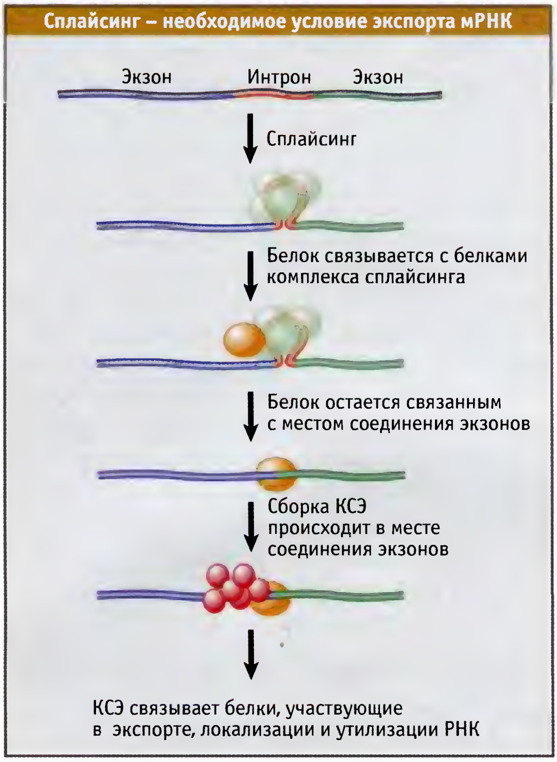


РИС. 26.16 Белки комплекса соединения экзонов (КСЭ) связываются с РНК, распознавая белки комплекса сплайсинга

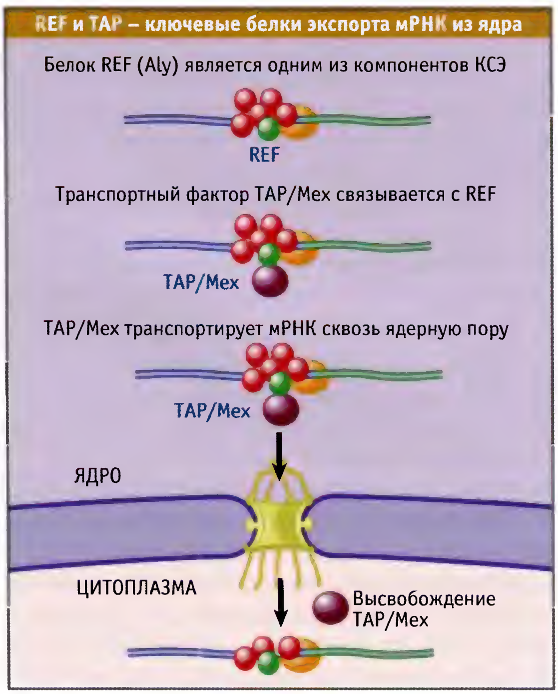


РИС. 26.17 Связываясь с фактором сплайсинга, белок REF остается в комплексе с РНК-продуктом. В таком состоянии он взаимодействует с белком ядерной поры, ответственным за экспорт РНК из ядра

белков. Первым с комплексом соединения экзонов контактирует один из факторов сплайсинга. После окончания сплайсинга комплекс соединения экзонов остается связан с мРНК в участке, расположенном непосредственно перед границей экзон-экзон. Комплекс соединения экзонов связывается исключительно со сплайсированной РНК и не взаимодействует с РНК, транскрибируемой с безынтронных генов.

Если удалить интрон из гена, РНК продукт такого гена будет экспортироваться в цитоплазму значительно медленнее. На основе этого можно предположить, что интрон служит сигналом для прикрепления аппарата экспорта к РНК. Сейчас мы можем рассматривать этот феномен в контексте серии белковых взаимодействий, как изображено на РИС. 26.17. В состав КСЭ входят белки, названные белками семейства REF (наиболее охарактеризованным членом этого семейства является белок Aly). Белки REF взаимодействуют с транспортным белком (называемым либо TAP, либо Mex), а тот в свою очередь — с ядерной порой.

Похожая система используется при идентификации сплайсированной РНК с помощью нонсенс мутаций в области, предшествующей последнему экзону. Такие мутации запускают деградацию РНК в цитоплазме (см. разд. 7.14 «Нонсенс-мутации запускают «систему надзора»).

26.11 Автосплайсинг интронов II группы происходит через формирование лассо

Ключевые понятия:

- Интроны группы II, обладая автокаталитической активностью, способны к самовырезанию из РНК.
- Границы сплайсинга и механизмы сплайсинга интронов группы II аналогичны таковым для ядерных интронов.
- Вторичная структура интрона группы II формирует каталитический сайт, напоминающий структуру ядерного интрона U6-U2.

Интроны в белок-кодирующих генах (фактически во всех генах, кроме ядерных генов, кодирующих тРНК) можно разделить на три основных класса. Интроны ядерной пре-мРНК идентифицируются только по наличию GU...AG динуклеотидов на 5′- и 3′-концах, а также сайта разветвления/пиримидинового региона рядом с 3′-концом. Они не демонстрируют никаких общих черт в особенностях своей вторичной структуры. Интроны групп I и II обнаруживаются в органеллах и бактериях. (Интроны группы I, кроме того, обнаруживаются в ядрах низших эукариот.) Интроны групп I и II классифицируются согласно своей внутренней организации. Каждому свойственен определенный тип вторичной структуры.

Интроны групп I и II обладают удивительной способностью к самовырезанию из РНК. Это явление носит

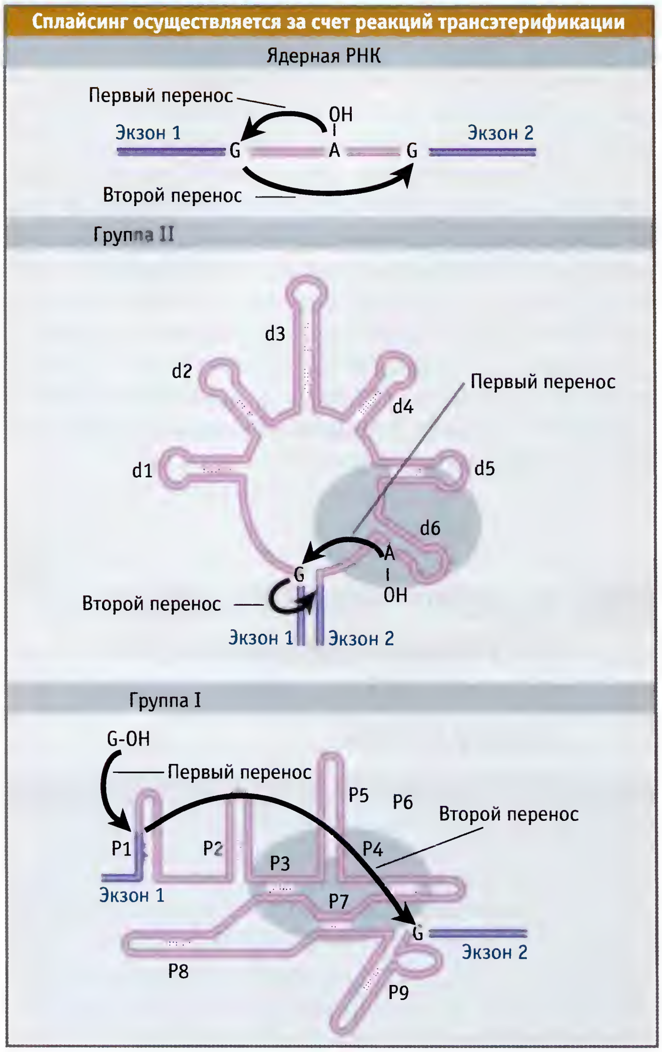


РИС. 26.18 Три типа реакций сплайсинга реализуются посредством двух реакций трансэтерификации. Во время первой реакции свободная ОН группа атакует место соединения экзона 1 и интрона. Во время второй – ОН группа, образовавшаяся на конце экзона 1 атакует место соединения интрона и экзона 2

название **автосплайсинг**. Интроны группы I более распространены, чем интроны группы II. Нет практически никакой связи между двумя этими классами интронов, однако и в том, и в другом случае РНК проявляет способность к аутосплайсингу *in vitro* без участия каких-либо белковых ферментативных активностей; тем не менее белки необходимы для корректного фолдинга *in vivo* (см. гл. 27 «Каталитическая РНК»).

На РИС. 26.18 показано, как в результате двух реакций трансэтерификации происходит вырезание интронов всех трех классов (ранее на рис. 26.6 это было показано для ядерных интронов). Во время первой реакции, 5′-граница интрон-экзон атакуется свободной гидроксильной группой (предоставляется внутренней 2′-ОН позицией в ядерных интронах и интронах группы II, и свободным гуанином в интронах группы I). Во время второй реакции свободная 3′-ОН-группа, расположенная на конце экзона, в свою очередь атакует 3′-границу интрон-экзон.

Можно провести параллели между интронами группы II и сплайсингом пре-мРНК. Вырезание митохондриальных интронов группы II происходит по тем же механизмам, что и сплайсинг ядерных пре-мРНК. И в том, и в другом случае за счет 5′–2′-связи формируется лассо. Пример лассо, образующегося при сплайсинге интронов группы II, приведен на РИС. 26.19. При инкубировании выделенной РНК с интронами группы II *in vitro*, реакция сплайсинга протекает в отсутствие каких-либо дополнительных компонентов. Это означает, что две реакции трансэтерификации, изображенные на рис. 26.18, могут осуществляться РНК, содержащей в своем составе интроны группы II, самостоятельно. Число фосфодиэфирных связей в ходе реакции не изменяется, и поэтому привнесения энергии извне для осуществления реакции не требуется; данная особенность может быть важным эволюционным приспособлением.

Во вторичной структуре интронов группы II выделяют несколько доменов, сформированных двухцепочечными стеблями и одноцепочечными петлями. Домен 5 отделен двумя основаниями от домена 6, в состав которого входит А, предоставляющий 2′-ОН группу для первой реакции трансэтерификации. Эти области составляют каталитический домен РНК. На РИС. 26.20 такая вторичная структура сравнивается со структурой, формируемой при спаривании U6 с U2 и U2 с сайтом разветвления. Сходство этих структур наталкивает на мысль о том, что U6 может играть каталитическую роль.

Особенности сплайсинга интронов группы II позволяют предположить, что сплайсинг развивался, начиная с автокаталитической реакции, осуществляемой

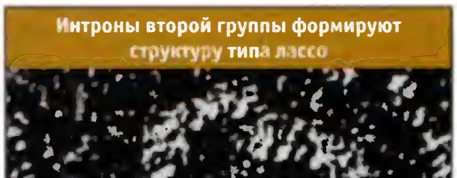


РИС. 26.19 Митохондриальный интрон второй группы, образовавший в ходе сплайсинга стабильную структуру типа лассо. Репродукция из Van der Veen, R., et al. *EMBO J.* 1987. 6: 1079-1084. Фотография любезно предоставлена leslie A. Grivell, European Molecular Biology Organisation

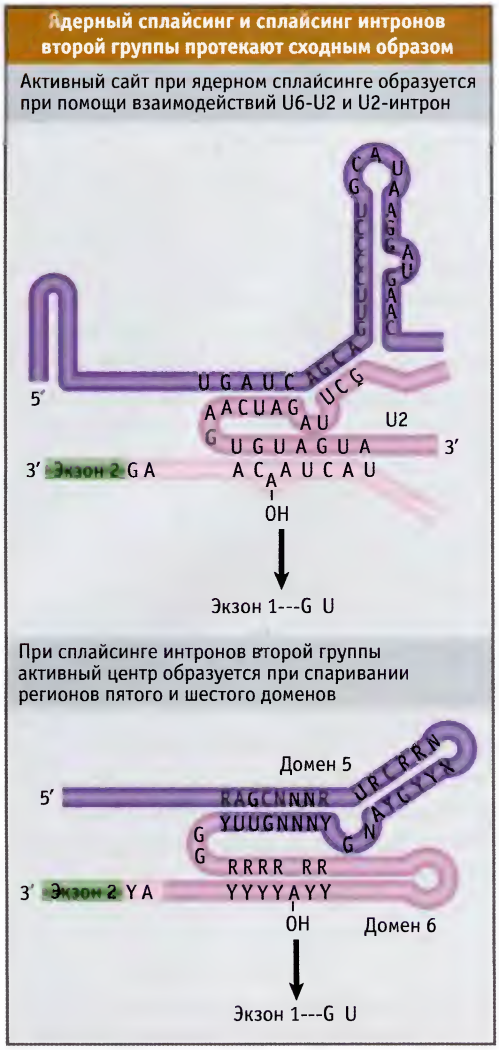


РИС. 26.20 При ядерном сплайсинге и сплайсинге интронов второй группы образуются похожие вторичные структуры. Последовательности, задействованные в ядерном сплайсинге, более специфичны. Для описания последовательностей, участвующих в сплайсинге интронов второй группы, достаточно указать их пуриновый/пиримидиновый состав

индивидуальной молекулой РНК и заключающейся в контролируемом вырезании внутренней последовательности. Вероятно, такая реакция требует от молекулы РНК определенной конформации или серии конформаций. Кроме того, она становится возможной только при *цис*-конформации РНК.

Своей способностью к самовырезанию в результате автокаталитического сплайсинга интроны группы II существенно отличаются от ядерных интронов, сплайсинг которых происходит при участии аппарата сплайсинга. Возможно, мяРНК сплайсосомы компенсируют неинформативность последовательности интрона, обеспечивая тем самым формирование определенной структуры РНК. Функции, выполняемые мяРНК, также могли эволюционировать от первоначальных автокаталитических систем. Реакция между мяРНК и субстратной пре-мРНК является *транс*-реакцией; можно представить, как способность U1 спариваться с 5′-сайтом сплайсинга, или U2 — с сайтом разветвления заменили похожую реакцию, которая осуществлялась аналогичными последовательностями, но расположенными в пределах одного интрона. Взаимодействия между мяРНК и субстратной пре-мРНК, а также мяРНК друг с другом пришли на смену серии конформационных изменений, которым подвергается РНК, сплайсируемая по механизму группы II. В результате субстратная РНК освободилась от необходимости нести в своем составе последовательности, способные обеспечить реакцию сплайсинга. По мере усложнения аппарата сплайсинга (и увеличения числа потенциальных субстратов), белки в этом процессе стали играть все более важную роль.

26.12 Альтернативный сплайсинг — это реализация различных вариантов сшивания экзонов между собой

Ключевые понятия:

- **Отдельные экзоны могут исключаться из РНК продукта или, наоборот, включаться в него в результате сшивания между собой разных пар сайтов сплайсинга.**
- **При использовании альтернативных вариантов сшивания экзонов длина конечного продукта сплайсинга может меняться.**
- **Детерминация пола у *Drosophila* является результатом серии реакций альтернативного сплайсинга в генах, кодирующих последовательные компоненты данного процесса.**
- **Р-элементы *Drosophila* подвергаются зародыш-специфичному альтернативному сплайсингу.**

Сплайсинг РНК, образующейся в результате транскрипции прерывистых генов, как правило, ведет к образованию единственного типа мРНК. Это свидетельствует о том, что реализуется лишь один вариант разрезания и сшивания интронов и экзонов. Однако РНК некоторых генов подвергается **альтернативному сплайсингу** — процессу, который обеспечивает кодирование одним геном нескольких последовательностей мРНК. В некоторых случаях окончательная модель экспрессии гена определяется первичным транскриптом, поскольку схема реакции сплайсинга изменяется в результате использования разных стартовых точек транскрипции или образования альтернативных 3′-концов. В других случаях реализуются разные варианты сплайсинга единственного первичного транскрипта, при этом внутренние экзоны заменяются на другие, добавляются или исключаются из состава зрелой мРНК. Иногда несколько продуктов сплайсинга единственного РНК-предшественника образуются в одной и той же клетке, иногда процесс альтернативного сплайсинга регулируется таким образом, что каждый конкретный вариант реализуется лишь в определенных условиях.

Самым острым вопросом в этой области остаются механизмы выбора того или иного варианта сплайсинга. Белки, участвующие в этом процессе, могут быть выявлены двумя способами. В некоторых модельных системах млекопитающих охарактеризовать альтернативный сплайсинг и выявить белки, необходимые для его осуществления, можно *in vitro*. У *D. melanogaster* нарушения процесса альтернативного сплайсинга могут быть вызваны мутациями либо в генах, РНК которых подвергается альтернативному сплайсингу, либо в генах, продукты которых необходимы для осуществления реакции сплайсинга.

На РИС. 26.21 приведены примеры альтернативного сплайсинга, в которых один сайт сплайсинга остается постоянным, а другой — варьирует. Большой Т-антиген и малый t-антиген вируса SV-40, а также продукты аденовирусного гена E1A образуются в результате сшивания разных 5′-сайтов с одним и тем же 3′-сайтом. В случае с T/t-антигенами 5′-сайт Т-антигена смещает терминирующий кодон t-антигена, в результате чего образуется больший по длине Т-антиген. При сплайсинге E1A транскриптов один из 5′-сайтов соединяется с последним экзоном согласно размерам разных рамок считывания, тем самым приводя к существенным изменениям С-концевой части белка. В данных примерах все описанные события сплайсинга происходят в каждой клетке, где эти гены экспрессируются, при этом образуются все возможные белковые продукты. Соотношение антигенов T/t в разных клетках может быть разным. Белок, выделенный из клеток, продуцирующих главным образом малые t-антигены, способен обеспечить предпочтительную продукцию малых (t)РНК экстрактами клеток других типов. Этот белок, названный ASF (фактор альтернативного сплайсинга), является аналогом фактора сплайсинга SF2, необходимого на ранних этапах сборки сплайсосомы и для первой реакции разрезания-сшивания (см. рис. 26.13). ASF/SF2 представляет собой РНК-связывающий белок, принадлежащий к SR-семейству. В случае когда мРНК имеет более одного 5′-сайта сплайсинга, предшествующих единственному 3′-сайту, повышенные концентрации ASF/SF2 способствуют использованию 5′-сайта, ближайшего к 3′-сайту в ущерб всем остальным сайтам. Эффект ASF/SF2 может нейтрализоваться действием другого фактора сплайсинга — SF5.

Точные молекулярные роли всех факторов, вовлеченных в данный процесс, до конца не выяснены. Однако мы можем заметить, что в механизмах альтернативного сплайсинга с использованием разных 5′-концов задействуются белки, участвующие в сборке сплайсосомы. В случае с T/t-антигенами эффект, возможно, основан на повышенном связывании белков SR с предпочтительным сайтом. Кроме того, альтернативный сплайсинг может быть результатом репрессии одного из сайтов. Активности экзонов 2 и 3 в гене тропонина Т мыши взаимоисключают друг друга; экзон 2 используется в тканях гладкой мускулатуры, экзон 3 — во всех остальных тканях. В клетках гладкой мускулатуры содержатся белки, которые связываются с повторяющимися элементами, расположенными на обоих концах экзона 3, и, тем самым, блокируют его 3′- и 5′-сайты.

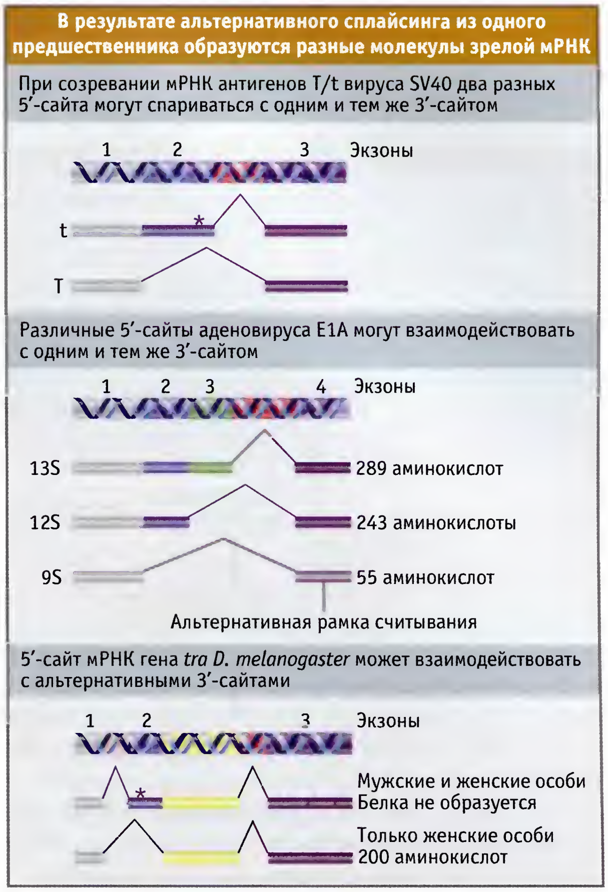


РИС. 26.21 **В результате альтернативного сплайсинга один ген становится способен кодировать образование разных белковых продуктов. Изменение сайтов сплайсинга может привести к появлению новых терминирующих кодонов (отмечены звездочками) или сдвигу рамки считывания**

Детерминация пола у *D. melanogaster* происходит в результате взаимодействия между собой серии генов, при этом события альтернативного сплайсинга определяют формирование мужских либо женских особей. Данный механизм представлен на РИС. 26.22. Отношение Х-хромосом к аутосомам определяется экспрессией гена *sxl*. Уровень экспрессии этого гена, в свою очередь, зависит от работы ряда других генов, последним из которых является ген *dsx*.

Процесс запускается в результате пол-специфичного сплайсинга гена *sxl*. Экзон 3 гена *sxl* содержит терминирующий кодон, который останавливает синтез функционального белка. Этот экзон входит в состав мРНК, свойственной мужским особям, и удален из мРНК женских особей. (Удаление экзона проиллюстрировано на другом примере на РИС. 26.23.) В результате, только у женских особей образуется белок Sxl. Содержание основных аминокислот в этом белке роднит его с другими РНК-связывающими пептидами.

В присутствии белка Sxl изменяется модель сплайсинга *transformer*-гена (*tra*). Как показано на рис. 26.21, в этот процесс вовлечены константный 5′-сайт и альтер-

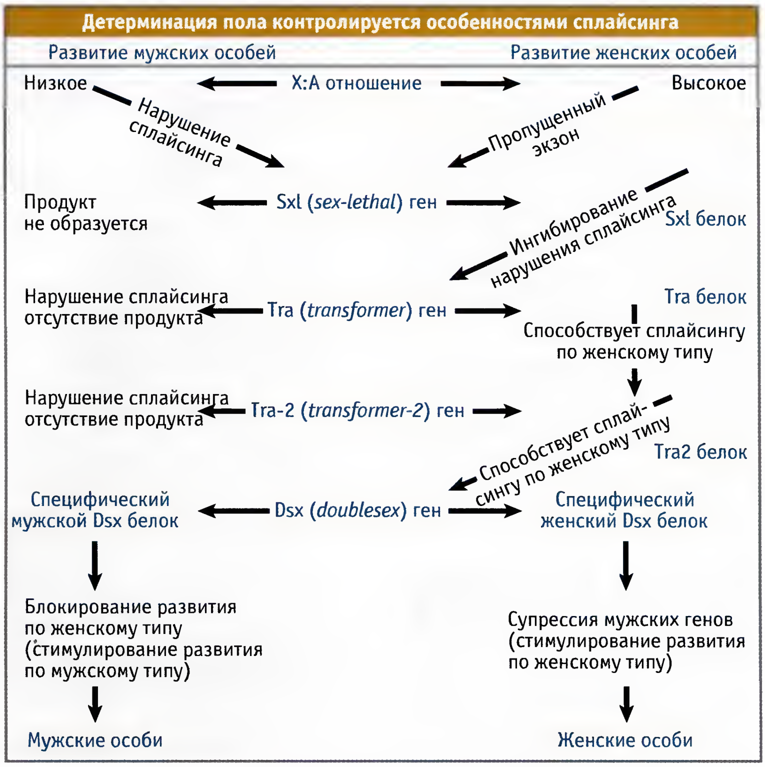


РИС. 26.22 Детерминация пола у *D. Melanogaster* является результатом событий альтернативного сплайсинга, происходящего у женских особей. Блокировка этого пути на любой из стадий приводит к развитию мужских особей

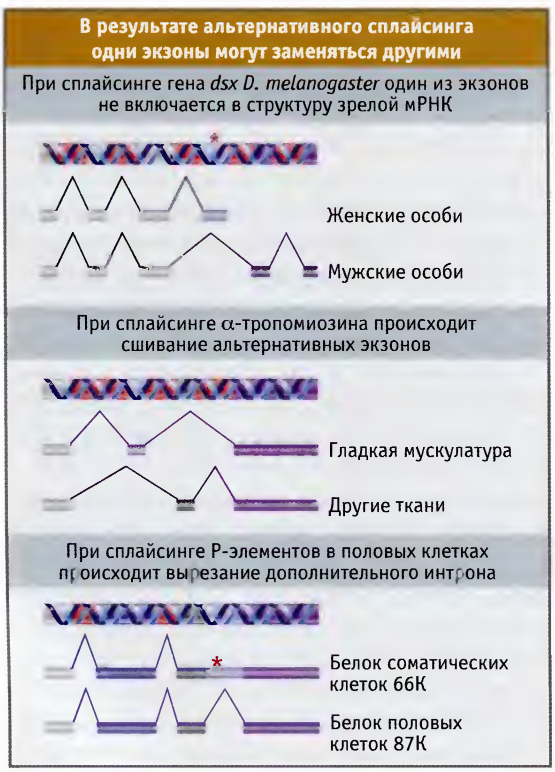


РИС. 26.23 События альтернативного сплайсинга, протекающие с участием обоих сайтов, могут привести к добавлению экзонов или замене одних экзонов на другие

нативные 3′-сайты. Одна из моделей сплайсинга свойственна как мужским, так и женским особям, в результате образуется РНК с ранним терминирующим кодоном. Белок Sxl ингибирует использование нормального 3′-сайта, связываясь с полипиримидиновым участком в его сайте разветвления. При пропуске одного 3′-сайта используется следующий 3′-сайт. Таким образом формируется кодирующая мРНК, специфичная для женских особей.

Ген *tra* кодирует образование белка только у женских особей; этот белок является регулятором сплайсинга. Ген *tra2* выполняет у женских особей похожую функцию (однако экспрессируется и у мужских зародышей). Белки Tra и Tra2 являются факторами сплайсинга семейства SR, действующими напрямую на транскрипты-мишени. Tra и Tra2 совместно (у женских особей) регулируют сплайсинг *dsx*.

На рис. 26.23 приведены примеры, когда сайты сплайсинга используются для добавления или замены экзонов или интронов. При этом так же образуются разные белковые продукты. В гене *doublesex (dsx)* женских особей сплайсинг происходит между 5′- и 3′-сайтами интрона 3, трансляция в этом случае терминируется в конце экзона 4. У мужских особей сплайсинг происходит между 5′-сайтом интрона 3 и 3′-сайтом интрона 4, минуя экзон 4, что делает возможной трансляцию экзонов 5 и 6. В результате альтернативного сплайсинга у каждого пола образуется свой набор белков: белковые продукты, свойственные мужским особям, блокируют дифферен-

цировку по женскому типу, а женские белки, в свою очередь, подавляют экспрессию мужских генов.

Альтернативный сплайсинг РНК гена *dsx* контролируется путем конкуренции между 3'-сайтами сплайсинга. РНК гена *dsx* содержит в своем составе элемент, расположенный после крайнего левого 3'-сайта сплайсинга и обладающий способностью связываться белком Tra2. Белки Tra и SR ассоциируют с Tra2 в этом сайте, который тем самым становится энхансером, способствующим связыванию U2AF с прилежащим пиримидиновым участком. В результате, формирование сплайсосомы у женских особей происходит с использованием именно этого 3'-сайта сплайсинга. Белки распознают энхансер совместно, ориентируясь как на определенную вторичную структуру региона, так и на нуклеотидную последовательность.

Процесс детерминации пола, таким образом, характеризуется определенной симметрией: начинаясь с событий сплайсинга, специфичных для мужских особей, и ведущих к образованию продуктов, из которых исключен экзон с терминирующим кодоном, он заканчивается событиями сплайсинга, специфичными для женских особей и ведущими к образованию продукта, содержащего в своем составе экзон с терминирующим кодоном. Все эти события имеют разную молекулярную основу. На первом контрольном этапе белок Sxl ингибирует стандартную модель сплайсинга. На последнем контрольном этапе совместное действие белков Tra и Tra2 обеспечивает течение сплайсинга по женскому типу.

При нормальном сплайсинге в белках Tra и Tra2 необходимости нет. В их отсутствие мухи развиваются в полноценные мужские особи. Tra и Tra2 не являются специфическими регуляторами реакции сплайсинга и этим отличаются от SF2, который необходим для осуществления основной реакции сплайсинга, однако может оказывать влияние и на выбор модели альтернативного сплайсинга.

Сплайсинг Р-элементов *D. melanogaster* демонстрирует тканеспецифичные черты. В соматических клетках происходят два события сплайсинга, а в половых в результате дополнительной стадии удаляется еще один интрон, содержащий в своем составе терминирующий кодон. При этом в половых клетках образуется более длинный белок (с другими свойствами). Механизм этого явления для контроля за транспозицией мы обсуждали в разд. 21.15 «Р-элементы могут быть активированы в клетках зародышевого пути», а сейчас следует отметить, что аппарат сплайсинга в этой системе имеет тканеспецифичный характер.

При инкубации РНК в экстракте, содержащем гетерологичные факторы сплайсинга (человека), сплайсинг пре-мРНК Р-элемента реализуется по пути, характерному для зародышевых клеток и сопровождающемуся удалением интрона 3. Экстракты соматических клеток *D. melanogaster*, однако, содержат белок, ингибирующий вырезание интрона. Белок связывается с последовательностью экзона 3. В отсутствие этой последовательности, интрон вырезается. Таким образом, функции белка сводятся к подавлению ассоциации сплайсосомы с 5'-сайтом интрона 3.

26.13 В реакциях *транс*-сплайсинга участвуют малые РНК

Ключевые понятия:

- Реакции сплайсинга, как правило, являются *цис*-реакциями, то есть происходят между участками одной молекулы РНК.
- *Транс*-сплайсинг, свойственный трипаносомам и червям, происходит, когда короткая последовательность (СЛ РНК) сплайсируется с 5'-концами разных предшественников мРНК.
- СЛ РНК имеет структуру, напоминающую структуру Sm-связывающего сайта U мяРНК, и, возможно, играет в реакции аналогичную роль.

И в механистическом, и в эволюционном контексте сплайсинг представляет собой интрамолекулярную реакцию, обеспечивающую контроль за вырезанием интронов на уровне РНК. В терминах генетики сплайсинг является *цис*-реакцией. Это означает, что *сплайсинг осуществляется только между последовательностями, расположенными на одной молекуле РНК.* В верхней части РИС. 26.24 показана классическая ситуация. Удаление интронов из молекулы РНК делает возможным сшивание экзонов этой же молекулы РНК. При этом *межмолекулярный* сплайсинг экзонов разных молекул РНК невозможен. Мы не можем утверждать, что *транс*-сплайсинг никогда не происходит между пре-мРНК транскриптов одного и того же гена. Однако если бы это происходило часто, экзоны одного гена обладали бы способностью комплементировать друг друга, и их нельзя было бы отнести к одной группе комплементации. Поэтому данное явление должно быть чрезвычайно редким.

Однако при создании определенных условий может происходить *транс*-сплайсинг. В примере, приведенном в нижней части рис. 26.24, в интроны двух молекул РНК были введены комплементарные последовательности. В результате спаривания между ними образуется Н-образная молекула. Эта молекула может подвергнуться *цис*-сплайсингу, при этом произойдет соединение экзонов, ковалентно связанных с интроном, а может подвергнуться *транс*-сплайсингу, при котором будут сшиваться экзоны расположенных рядом молекул РНК. Обе эти реакции протекают *in vitro*.

Другой ситуацией, в которой становится возможен транс-сплайсинг, является ситуация, при которой одна субстратная РНК предоставляет 5'-сайт, а другая — 3'-сайт сплайсинга вместе с соответствующими последовательностями, расположенными за сайтами (к примеру, следующий 5'-сайт сплайсинга или энхансер сплайсинга). Механизм данного процесса напоминает механизм сплайсинга с распознаванием экзона (см. правую часть рис. 26.12) и не требует (для сплайсинга *in vitro*), чтобы левый и правый сайты сплайсинга располагались на одной и той же молекуле РНК.

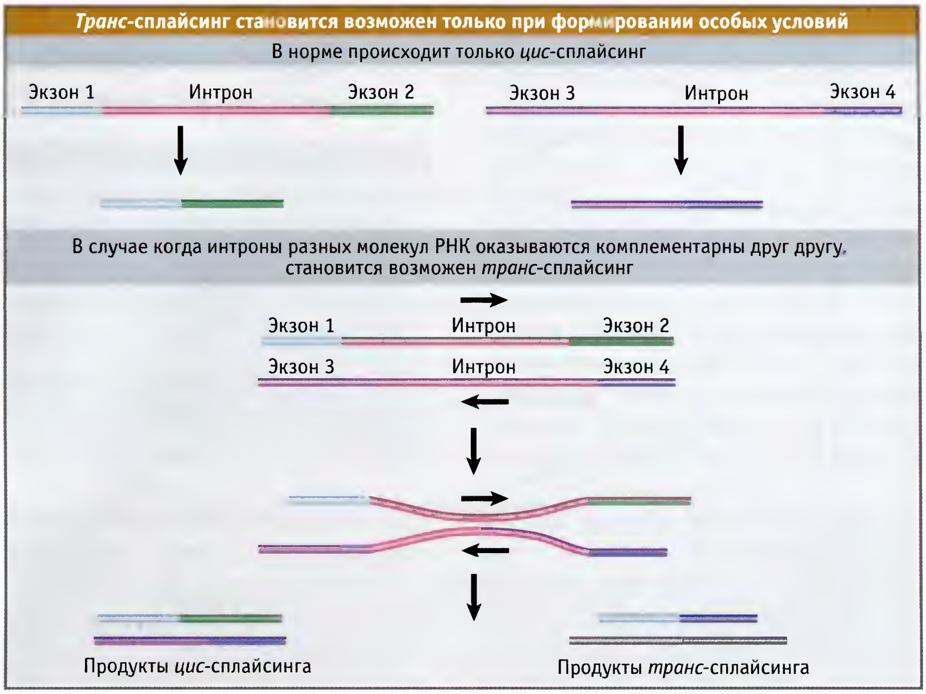


РИС. 26.24 Как правило, сплайсинг осуществляется между экзонами, находящимися в *цис*-положении, то есть расположенными на одной молекуле РНК. В некоторых случаях при формировании связей между интронами разных молекул РНК становится возможен *транс*-сплайсинг

Полученные результаты показывают, что никаких *механических* помех для осуществления *транс*-сплайсинга нет. Кроме того, они исключают модель сплайсинга, согласно которой сплайсосома последовательно движется вдоль молекулы РНК. Сплайсосома способна распознавать 5′- и 3′-сайты разных молекул РНК, расположенных в непосредственной близости друг от друга.

Хотя явление *транс*-сплайсинга является довольно редким, в определенных ситуациях оно наблюдается *in vivo*. Одним из примеров транс-сплайсинга может служить наличие одинаковой 35-нуклеотидной последовательности на 5′-конце многих мРНК у трипаносом. При этом каждой отдельной транскрипционной единицей лидерная последовательность не кодируется. Вместо этого она транскрибируется с повторяющейся единицы, расположенной по всему геному, в отдельную молекулу РНК, несущую дополнительную последовательность на 3′-конце. Как показано на РИС. 26.25, эта РНК несет 35-нуклеотидную лидерную последовательность, за которой следует последовательность 5′-сайта сплайсинга. Последовательности мРНК содержат 3′-сайт сплайсинга, предшествующий в зрелой мРНК лидерной последовательности.

После соединения лидерной последовательности и мРНК в результате реакции *транс*-сплайсинга, 3′-регион первой и 5′-регион второй образуют вместе 5′- и

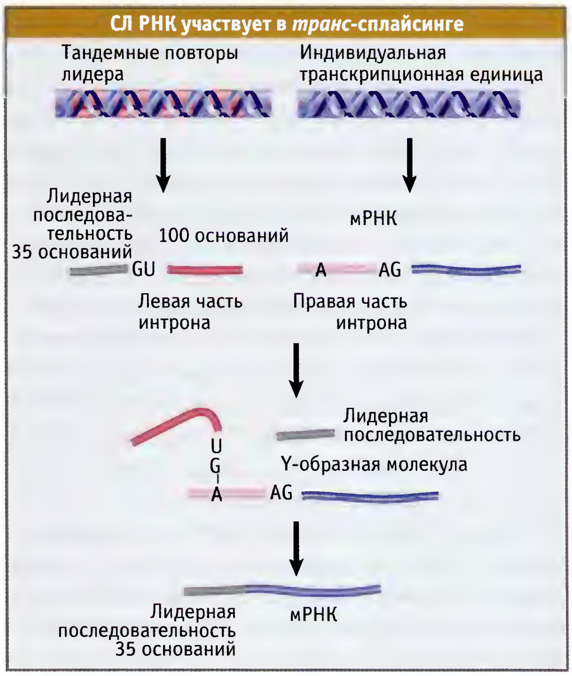


РИС. 26.25 СЛ РНК представляет собой экзон, связывающийся с первым экзоном мРНК при *транс*-сплайсинге. Реакция опосредована такими же взаимодействиями, что и ядерный *цис*-сплайсинг, только вместо структуры типа лассо образуется Y-образная структура

3′-половины интрона. В ходе сплайсинга 5′–2′ связь формируется в результате обычной реакции между динуклеотидом GU 5′-интрона и последовательностью сайта разветвления рядом с динуклеотидом AG 3′-интрона. Две части интрона не связаны между собой ковалентной связью, потому вместо лассо образуется Y-образная молекула.

Похожая ситуация наблюдается при анализе экспрессии генов актина у *Caenorhabditis elegans*. Три актиновых мРНК (и ряд других РНК) имеют одинаковую 22-нуклеотидную лидерную последовательность на 5′-конце. Лидерная последовательность не кодируется геном актина, а транскрибируется независимо в составе 100-нуклеотидной РНК с гена, расположенного в разных участках генома. Кроме того, *транс*-сплайсинг наблюдается в хлоропластах.

РНК, предоставляющая 5′-экзон для *транс*-сплайсинга, называется **СЛ РНК** (РНК со сплайсированным лидером). СЛ РНК, обнаруженные у некоторых видов трипаносом и нематод (*C. elegans*), демонстрируют общие черты. Они имеют одинаковую вторичную структуру, представленную тремя элементами типа стебель-петля и одноцепочечным регионом, напоминающим Sm-связывающий сайт. СЛ РНК существуют в виде мяРНП, относящихся к классу Sm мяРНП. У трипаносом обнаружены U2, U4 и U6 мяРНК и отсутствуют U1 и U5 мяРНК. Отсутствие U1 мяРНК может быть объяснено свойствами СЛ РНК, которая способна выполнять функции U1 мяРНК в 5′-сайте сплайсинга. СЛ РНК содержит в своем составе последовательность, способную обеспечить в сайте экзон-интрон протекание всех тех процессов, за которые отвечает U1 мяРНК.

У *C. elegans* обнаружено два типа СЛ РНК. СЛ1 РНК (была открыта первой) присоединяется только к тем кодирующим последовательностям, которым предшествовали 5′-нетранслируемые регионы (наиболее общая ситуация). СЛ2 РНК используется в тех случаях, когда пре-мРНК содержит две кодирующие последовательности; СЛ2 РНК присоединяется ко второй последовательности, тем самым приводя к высвобождению первой и превращению ее в независимую молекулу РНК. Около 15% генов *C. elegans* организованы в транскрипционные единицы, которые содержат более одного гена (чаще 2–3). Значение такой формы организации для контроля за экспрессией генов не понятно. Эти транскрипционные единицы не похожи на опероны, в которые гены объединены согласно их функциональному параметру.

Реакции *транс*-сплайсинга с участием СЛ РНК могут представлять собой очередной шаг в эволюции аппарата сплайсинга пре-мРНК. СЛ РНК способна распознавать 5′-сайт сплайсинга, что, возможно, зависит от особой конформации РНК. Оставшиеся функции, необходимые для осуществления сплайсинга, обеспечиваются независимыми мяРНП. СЛ РНК может функционировать без участия белков, чем напоминает U1 мяРНП, распознавание которой 5′-сайта зависит только от свойств РНК.

26.14 Сплайсинг тРНК дрожжей идет через разрезание и сшивание предшественника

Ключевые понятия:

- Сплайсинг тРНК представляет собой серию реакций разрезания и лигирования

Процесс сплайсинга зависит от коротких консенсусных последовательностей и представляет собой серию реакций трансэтерификации, в которых этапы разрыва и образования связей координированы. Существуют разные механизмы сплайсинга тРНК, однако в основе всех них лежат отдельные реакции расщепления и лигирования молекул.

59 из 272 ядерных генов тРНК у дрожжей *S. cerevisiae* имеют прерывистое строение. В каждом случае в состав гена входит один интрон, расположенный на расстоянии одного нуклеотида от 3′-конца антикодона. Длина интрона варьирует от 14 до 60 пн. Нуклеотидные последовательности интронов генов, кодирующих тРНК для одной аминокислоты, сходны между собой, в то время как интроны генов, кодирующих тРНК для разных аминокислот, не похожи друг на друга. *Нет никакой канонической последовательности, которая могла бы узнаваться ферментами сплайсинга.* Это справедливо и для прерывистых генов тРНК растений, амфибий и млекопитающих.

Все интроны содержат последовательность, комплементарную антикодону тРНК. Это приводит к формированию альтернативной конформации антикодоновой петли, при которой ее длина увеличивается за счет спаривания оснований антикодона с интроном. Пример такого взаимодействия приведен на РИС. 26.26. Меняется только антикодоновая петля, остальная часть молекулы сохраняет свою обычную структуру.



РИС. 26.26 Интрон тРНК^{Phe} дрожжей спаривается с антикодоном, изменяя структуру антикодонового плеча молекулы. Взаимодействие между выделенным основанием, расположенным в стебле, и петлей, формируемой интроном в молекуле предшественника, может являться необходимым условием сплайсинга

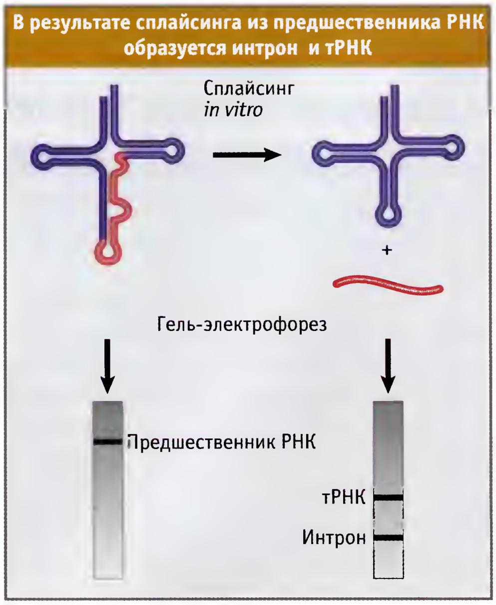


РИС. 26.27 Сплайсинг тРНК дрожжей *in vitro* можно увидеть с помощью анализа предшественника РНК и продуктов сплайсинга методом гель-электрофореза

Точная нуклеотидная последовательность и размер интрона при этом значения не имеют. Большинство мутаций, возникающих в интроне, на сплайсинг не влияют. *В основе сплайсинга тРНК лежит узнавание вторичной структуры тРНК, а не нуклеотидной последовательности иинтрона.* В формировании соответствующей вторичной структуры участвуют регионы, расположенные в разных частях молекулы, включая участок между акцепторным участком и D-плечом, участок в Tψ C-плеча и в особенности антикодоновое плечо. Все эти взаимодействия напоминают формирование структуры тРНК, необходимой для осуществления синтеза белка (см. гл. 8 «Синтез белков»).

Тем не менее структура интрона все же имеет значение. Спаривание между основаниями петли интрона и одноцепочечным стеблем является необходимым условием сплайсинга. Мутации в других положениях, влияющие на это взаимодействие (к примеру, приводящие к изменению схемы спаривания), также влияют на сплайсинг. Правила, описывающие пригодность тРНК к сплайсингу, напоминают правила, описывающие процесс распознавания тРНК аминоацил-тРНК синтетазой (см. разд. 9.9 «Синтетазы присоединяют аминокислоты к тРНК»).

У температуро-чувствительных мутантов дрожжей с нарушенным механизмом удаления интронов предшественники тРНК, имеющие прерывистое строение, накапливаются в ядре. Такие молекулы могут быть использованы в качестве субстрата в бесклеточной системе, полученной из клеток дикого типа. За сплайсингом предшественника можно легко проследить, определяя степень уменьшения размеров образовавшегося продукта. Об этом свидетельствует изменение положения полосы РНК при гель-электрофорезе, как показано на РИС. 26.27. Уменьшение размеров предшественника коррелирует с появлением на геле полосы, соответствующей интрону.

Бесклеточный экстракт может быть фракционирован путем анализа его способности к сплайсингу тРНК. Реакция *in vitro* требует участия АТФ. Анализ роли АТФ в реакции позволил разбить реакцию *на два отдельных этапа, каждый из которых катализируется своим ферментом.*

- Первый этап не требует участия АТФ. Он представляет собой расщепление фосфодиэфирной связи, протекающее как необычная нуклеазная реакция, катализируемая эндонуклеазой.
- На втором этапе происходит образование связи, требующее присутствия АТФ. Эту реакцию называют лигированием, а фермент, ее осуществляющий, — **РНК-лигазой**.

26.15 Эндонуклеаза сплайсинга распознает тРНК

Ключевые понятия:

- Эндонуклеаза расщепляет предшественник тРНК по обоим концам интрона.
- Эндонуклеаза дрожжей представляет собой гетеротетрамер с двумя (родственными) каталитическими субъединицами.
- Для распознавания сайтов расщепления эндонуклеаза использует механизмы, позволяющие оценить положение фермента относительно определенного участка тРНК.
- Нуклеаза архей имеет более простую структуру и распознает мотив «изгиб-спираль-изгиб» в субстрате.

За специфическое распознавание интрона ответственна эндонуклеаза. Она расщепляет молекулу предшественника по обоим концам интрона. Эндонуклеаза дрожжей представляет собой гетеротетрамерный белок. Ее активности приведены на РИС. 26.28. Родственные субъединицы Sen34 и Sen2 расщепляют предшественник в 3′- и 5′-сайтах соответственно. Субъединица Sen54 способна обеспечить распознавание сайтов расщепления путем измерения расстояния от некоторой точки в структуре тРНК. Эта точка располагается в изгибе зрелой L-образной структуры тРНК. Роль субъединицы Sen15 не выяснена, однако установлено, что кодирующий ее ген необходим для нормальной жизнедеятельности дрожжей. Спаривание между первым основанием антикодоновой петли и основанием, предшествующим 3′-сайту сплайсинга, является условием расщепления в 3′-сайте.

Интересные данные об эволюции сплайсинга тРНК можно получить, анализируя эндонуклеазу архей. Этот фермент представляет собой гомодимер или гомотетрамер, в котором каждая субъединица имеет активный центр (несмотря на то что в тетрамере функционируют только два из них), катализирующий расщепление в соответствующих сайтах сплайсинга. В составе субъединиц фермента архей выделяют последовательности,

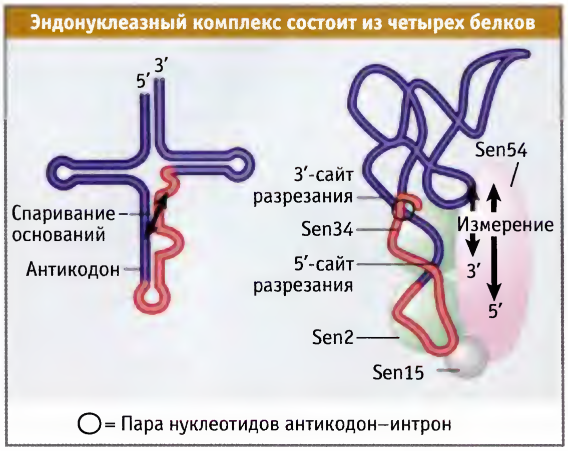


РИС. 26.23 Реакции расщепления в 3′- и 5′-сайтах сплайсинга пре-тРНК у *S. cerevisiae* катализируются разными субъединицами эндонуклеазы. Еще одна субъединица отвечает за определение местоположения сайтов разрезания, оценивая расстояния в зрелой молекуле

схожие с последовательностями активных центров субъединиц Sen34 и Sen2 дрожжей. Фермент архей распознает свой субстрат способом, отличным от способа работы дрожжевого фермента. Вместо измерения расстояния от определенных последовательностей он узнает структурную особенность молекулы предшественника тРНК, названную мотивом «изгиб-спираль-изгиб». Как показано на РИС. 26.29, расщепление происходит в области двух изгибов.

Эволюционное происхождение сплайсинга тРНК предшествовало разделению архей и эукариот. Если он возник в результате встраивания интронов в тРНК, это событие должно было быть очень древним.

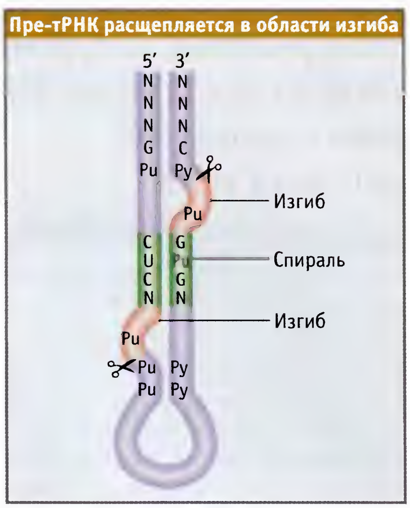


РИС. 26.29 При сплайсинге тРНК архей эндонуклеаза расщепляет каждую цепь в области изгиба мотива «изгиб-спираль-изгиб»

26.16 Расщепление и лигирование тРНК представляют собой две отдельные реакции

Ключевые понятия:

- В результате вырезания интрона образуются две половины молекулы тРНК, которые, соединяясь друг с другом, формируют зрелую структуру
- Эти половины имеют необычные концы: на 5′-конце — гидроксильная группа, а на 3′-конце — 2′-3′ циклический фосфат.
- 5′-ОН-группа фосфорилируется полинуклеотидкиназой, циклическая фосфатная группа под действием фосфодиэстеразы раскрывается и образует 2′-фосфатный конец и 3′-ОН-группу, концы экзонов соединяются РНК-лигазой, а 2′-фосфат удаляется фосфатазой.

Суммарно все процессы, составляющие реакцию сплайсинга, представлены на РИС. 26.30. Продуктами расщепления являются линейный интрон и две половины молекулы тРНК. Эти промежуточные продукты имеют уникальные концы. На каждом 5′-конце располагается гидроксильная группа; на каждом 3′-конце — 2′–3′-циклический фосфат. (Все другие известные ферменты сплайсинга РНК расщепляют фосфатную связь с другой стороны.)

Основания двух половин тРНК спариваются с образованием тРНК-подобной структуры. При добавлении АТФ осуществляется вторая реакция, в ходе которой изменяются оба необычных конца, образовавшиеся под действием эндонуклеазы.

Циклическая фосфатная группа открывается с образованием 2′-фосфатного конца. Эта реакция происходит под действием фосфодиэстеразы. Продукт реакции имеет 2′-фосфатную группу и 3′-ОН-группу.

5′-ОН-группа, образовавшаяся под действием нуклеазы, должна превратиться в 5′-фосфатную группу в результате фосфорилирования. При этом формируется сайт, в котором 3′-ОН-группа и 5′-фосфатная группа располагаются рядом. Ковалентная целостность полинуклеотидной цепи восстанавливается лигазой.

Все три ферментативные активности — фосфодиэстеразная, полинуклеотидкиназная и аденилатсинтазная (обеспечивает выполнение лигазной функции) — организованы в разные функциональные домены на одном и том же белке. Действуя последовательно, они осуществляют соединение двух половин молекулы тРНК.

После сплайсинга молекула становится непрерывной с 5′–3′-фосфатной связью в сайте сплайсинга, но она также содержит 2′-фосфатную группу — маркер осуществления сплайсинга. Эта лишняя группа должна быть удалена фосфатазой.

2′–3′-циклический фосфат возникает при сплайсинге тРНК у растений и млекопитающих. Было установлено, что механизм его образования у растений аналогичен таковому у дрожжей, однако отличается у млекопитающих.

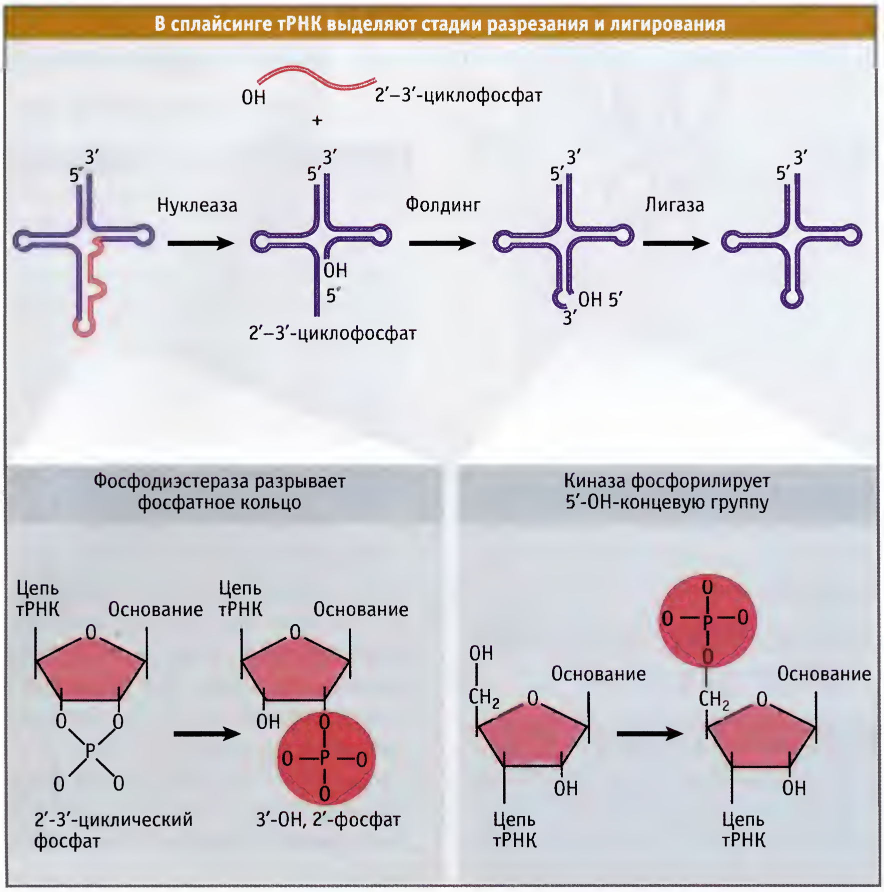


РИС. 26.30 Сплайсинг тРНК требует участия ферментов с нуклеазной и лигазной активностью. Нуклеаза расщепляет молекулу РНК на границе интрон-экзон, при этом образуется 2′–3′-циклический фосфат и концевая 5′-OH-группа. Циклический фосфат разрывается с образованием 3′-OH- и 2′-фосфатной группы. 5′-OH-группа фосфорилируется. После высвобождения интрона два фрагмента молекулы тРНК сворачиваются в тРНК-подобную структуру, имеющую свободные 3′-OH- и 5′-фосфатной группы на концах. Этот разрыв сшивается лигазой

Сплайсинг предшественников дрожжевой тРНК может осуществляться также в экстракте, полученном из зародышевых пузырьков (ядер) ооцитов *Xenopus*. Из этого следует, что реакция не является видоспецифичной. У *Xenopus*, по всей вероятности, имеются ферменты, способные узнавать интроны тРНК дрожжей.

Способность к сплайсингу продуктов генов, кодирующих тРНК, является высококонсервативной и, по-видимому, имеет иное происхождение в сравнении с другими реакциями сплайсинга (например, сплайсинга ядерной пре-мРНК). Расщепление и образование связей при сплайсинге тРНК контролируются последовательностями, расположенными за пределами интрона. В ходе других реакций сплайсинга используется трансэтерификация, в ходе которой перенос связей происходит напрямую, а все последовательности, необходимые для осуществления реакции, лежат внутри интрона.

26.17 Каскад реакций, запускаемых в ответ на накопление белков, связан с процессом сплайсинга тРНК

Ключевые понятия:

- Ire1p — белок внутренней ядерной мембраны, N-концевой домен этого белка обращен в просвет ЭПР, а C-концевой домен — внутрь ядра.
- Связывание несвернувшегося белка с N-терминальным доменом запускает реакцию автофосфорилирования, в результате чего активируется C-терминальная нуклеаза.
- Активированная нуклеаза расщепляет мРНК Hac1, при этом интрон высвобождается, а экзоны лигируются тРНК-лигазой.
- Сплайсированная мРНК Hac1 кодирует транскрипционный фактор, активирующий гены шаперонов, способствующих фолдингу белков.

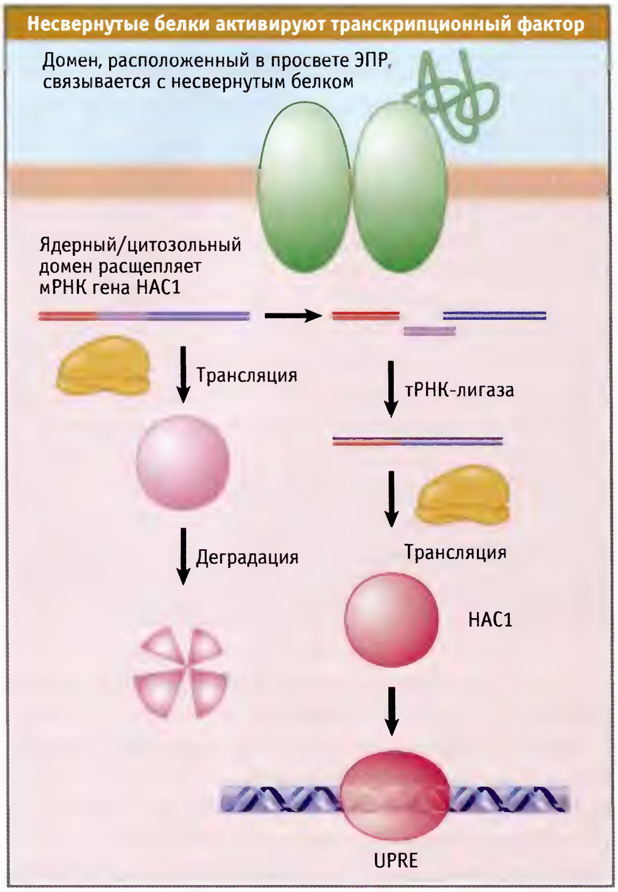


РИС. 26.31 **Каскад реакций, запускающихся в ответ на накопление в просвете ЭР не свернувшихся в правильную конформацию белков, обеспечивается альтернативным сплайсингом мРНК гена НАС1, в результате которого образуется транскрипционный фактор, распознающий UPRE**

Необычная система сплайсинга, связанная со сплайсингом тРНК, опосредует ответ дрожжевой клетки на накопление несвернувшихся в правильную конформацию белков. Накопление таких полипептидов в просвете эндоплазматического ретикулума (ЭПР) запускает каскад реакций, в ходе которых активируется транскрипция генов шаперонов, способствующих фолдингу белков в ЭПР. При этом сигнал должен быть передан из просвета ЭПР в ядро.

Активатором данного процесса служит белок Ire1p. Это интегральная мембранная протеинкиназа (Ser/Thr), концевые домены которой расположены с обеих сторон мембраны ЭПР. N-концевой домен в просвете ЭПР реагирует на появление несвернувшихся белков, предположительно при их связывании с его выступающим за пределы мембраны мотивом. Это приводит к агрегации мономеров, запуску реакции автофосфорилирования и активации С-концевого домена на другой стороне мембраны.

Гены, активируемые в результате этого каскада, имеют общий промоторный элемент — элемент ответа на развернутый белок. Транскрипционный фактор Hac1p, связывающийся с UPRE, образуется в ответ на накопление несвернувшихся белков. Запуск продукции Hac1p происходит в результате взаимодействия Ire1p с мРНК Hac1.

Все стадии процесса приведены на РИС. 26.31. В нормальных условиях, когда данный каскад не активирован, мРНК Hac1 транслируется в белок, который быстро деградирует. Активированный Ire1p влияет на сплайсинг мРНК Hac1 таким образом, что в результате трансляции этой РНК образуется более стабильный белок, который служит транскрипционным фактором, активирующим гены с UPRE.

В эту реакцию вовлечены необычные компоненты сплайсинга. Ire1p обладает эндонуклеазной активностью и действует напрямую на мРНК Hac1, расщепляя ее в области границ сплайсинга. Затем концы образовавшихся фрагментов лигируются тРНК-лигазой, обеспечивающей сплайсинг тРНК. Эндонуклеазная реакция напоминает реакцию расщепления тРНК во время сплайсинга.

Где же происходят модификации мРНК Hac1? Ire1p, вероятно, локализуется во внутренней ядерной мембране. N-концевой сенсорный домен молекулы обращен в просвет ЭПР, а С-концевой киназно/нуклеазный домен — в ядро. Такая ориентация дает возможность белку действовать на мРНК Hac1 напрямую до того, как она будет транспортирована в цитоплазму. Кроме того, это обеспечивает свободный доступ тРНК лигазы к субстрату. Никакого очевидного сходства между природой нуклеазной активности Ire1p и эндонуклеазы сплайсинга тРНК не наблюдается, в связи с чем не ясно, каким образом эволюционировали эти специализированные системы.

26.18 3′-концы транскриптов polI и polIII образуются в результате терминации транскрипции

Ключевые понятия:

- РНК-полимераза I терминирует транскрипцию, достигая 18-нуклеотидной терминаторной последовательности.
- РНК-полимераза III терминирует транскрипцию, достигая поли$(U)_4$-последовательности, расположенной внутри G-C-богатого региона.

3′-конец молекулы РНК может быть образован двумя способами. Некоторые РНК-полимеразы терминируют транскрипцию, достигнув определенной (терминаторной) последовательности в ДНК, как показано на РИС. 26.32. Для других РНК-полимераз не характерна дискретная терминация, и они продолжают транскрипцию до тех пор, пока 3′-конец не будет сформирован в результате расщепления молекулы РНК эндонуклеазой. Пример работы такого фермента приведен на РИС.

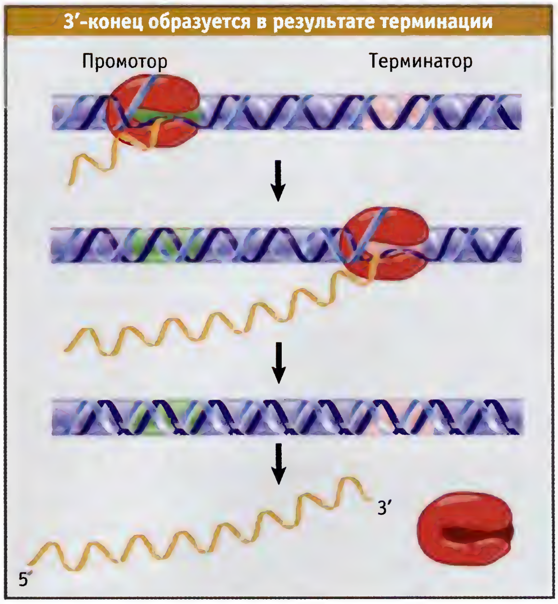


РИС. 26.32 В случае когда 3′-конец молекулы образуется в результате терминации, РНК-полимераза и РНК-продукт высвобождаются из транскрипционного комплекса в определенном участке последовательности ДНК (терминаторе)

Данных о механизмах терминации эукариотических РНК-полимераз существенно меньше, чем данных об инициации транскрипции. РНК-полимеразам I и III свойственна дискретная терминация (как и бактериальным полимеразам), способ терминации, используемый РНК-полимеразаой II, не выяснен.

РНК-полимеразой I синтезирует транскрипты генов основных рРНК. Молекула предшественника имеет большой размер и в дальнейшем подвергается обширному процессингу. Терминация происходит в сайте, расположенном примерно через 1000 оснований после участка, соответствующего 3′-концу зрелой молекулы рРНК. Формирование 3′-конца осуществляется путем расщепления молекулы предшественника. Терминация происходит при распознавании вспомогательными факторами 18-нуклеотидной терминаторной последовательности.

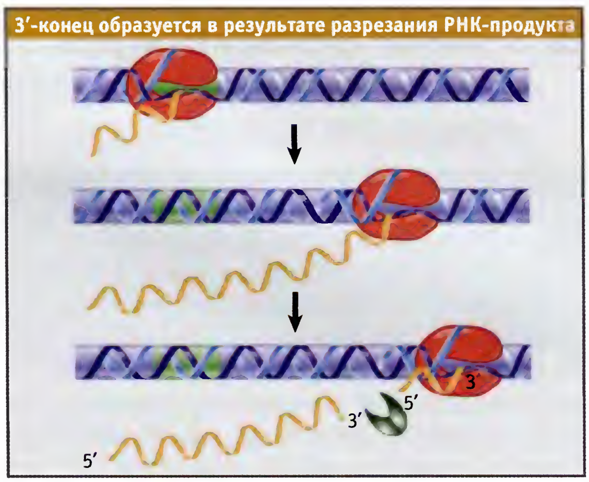


РИС. 26.33 В случае когда 3′-конец образуется в результате разрезания молекулы РНК-продукта, РНК-полимераза продолжает транскрипцию до тех пор, пока эндонуклеаза не разрежет вновь ситезированную РНК в определенном сайте

РНК полимеразой III *in vitro* образует молекулы с такими же 5′- и 3′-концами, что и у молекул, синтезируемых *in vivo*. Механизм реакции терминации напоминает таковой для реакции терминации, осуществляемой бактериальной РНК-полимеразой в ρ-независимом терминаторе (см. разд. 11.21 «Два типа терминаторов *E. coli*»). Терминация, как правило, происходит на втором U в последовательности из четырех остатков U, однако обнаруживаются транскрипты, заканчивающиеся тремя или даже четырьмя U. Такая гетерогенность свойственна молекулам, синтезированным *in vivo*, демонстрируя, что данная особенность является индивидуальным свойством каждой отдельной реакции терминации.

Как и прокариотические терминаторы, последовательность поли$(U)_4$ входит в состав G–C-богатого региона. Несмотря на наличие последовательностей с двусторонней симметрией, эта особенность не имеет для терминации никакого значения. Мутации, нарушающие симметрию, не препятствуют нормальному синтезу РНК. В терминации не участвуют последовательности, расположенные после поли$(U)_4$-участка. Все они могут быть заменены без каких-либо последствий для терминации.

Самой по себе поли$(U)_4$-последовательности недостаточно для терминации, о чем свидельствует наличие повторов из четырех оснований U внутри транскрипционной единицы. (Однако внутренних повторов U_5, обладающих большей терминирующей способностью, чем U_4, не обнаруживается, в то время как терминаторы U_5 встречаются чаще.) Принципиальным является распознавание U_4 последовательности в составе G–C региона.

Как же осуществляется реакция терминации? В ее основе не могут лежать слабые взаимодействия rU-dA в РНК-ДНК гибридном регионе на конце транскрипта, поскольку чаще всего транскрибируются только первые два основания U. Возможно, G-C-богатый участок обеспечивает замедление хода фермента, однако вряд ли он является аналогом шпильки, участвующей в прокариотической транскрипции. Какими механизмами обеспечивается столь специфичная реакция фермента на такую короткую последовательность, не выяснено. В противоположность процессу инициации транскрипции, который РНК полимераза III не может осуществлять самостоятельно, способность терминировать транскрипцию, по всей видимости, является свойством самого фермента.

26.19 3′-концы мРНК образуются в результате расщепления и полиаденилирования первичного транскрипта

Ключевые понятия:

- Последовательность AAUAAA является маркером 3′-конца, образующегося в результате расщепления молекулы мРНК с последующим полиаденилированием.
- Реакция происходит при участии белкового комплекса, в состав которого входят фактор специфичности, эндонуклеаза и поли(A)-полимераза.
- Фактор специфичности и эндонуклеаза расщепляют РНК в участке, расположенном после последовательности AAUAAA.
- Фактор специфичности и поли(A)-полимераза добавляют примерно 200 остатков A к формируемому 3′-концу.
- A-U-богатая последовательность на 3′-конце мРНК контролирует процессы поли- и деаденилирования в цитоплазме эмбрионов *Xenopus*.

3′-конец мРНК образуется в результате разрезания синтезируемой цепи РНК, за которым следует полиаденилирование новообразованного 3′-конца. Присоединению поли(A) к ядерной РНК может помешать аналог дезоксиаденозинтрифосфорной кислоты — 3′-дезоксиаденозин, известный также под названием **кордицепин**. Несмотря на то что кордицепин не препятствует синтезу ядерной РНК, его добавление нарушает выход мРНК в цитоплазму. Это свидетельствует о *необходимости* полиаденилирования для образования зрелой мРНК из ядерной пре-РНК.

Процесс образования 3′-конца представлен на РИС. 26.34. При достижении РНК-полимеразой в ходе транскрипции сайта, соответствующего 3′-концу транскрипта, и узнавании последовательностей-мишеней для действия эндонуклеазы происходит расщепление синтезируемой молекулы РНК и полиаденилирование ее 3′-конца. И разрезание РНК, и полиаденилирование осуществляются единым комплексом процессинга. Полиаденилирование стабилизирует молекулу мРНК, защищая ее от деградации с 3′-конца. К этому времени 5′-конец молекулы уже стабилизирован в результате кэпирования. После разрезания синтезируемой цепи РНК-полимераза продолжает транскрипцию, образовавшийся при этом 5′-конец остается незащищенным.

Разрезание транскрипта косвенно запускает механизм терминации транскрипции РНК-полимеразой II. Экзонуклеаза связывается с 5′-концом молекулы РНК, синтез которой продолжается и после разрезания. Фермент деградирует РНК быстрее, чем она синтезируется, до тех пор, пока не догонит РНК-полимеразу. Вспомогательные белки, связывающиеся с карбокси-терминальным доменом полимеразы, способствуют отделению РНК-полимеразы от ДНК и терминации транскрипции. В целом механизм процесса напоминает таковой для ρ-зависимой терминации, свойственной бактериальной РНК-полимеразе (см. разд. 11.22 «Как работает ρ-фактор?»). Это объясняет, почему места локализации сайтов терминации РНК-полимеразы II не определяются однозначно, а могут варьировать внутри протяженного региона, расположенного после сайта, соответствующего 3′-концу мРНК.

Общей чертой всех мРНК высших эукариот (кроме дрожжей) является наличие высококонсервативной последовательности AAUAAA в участке, расположенном на 11–30 нуклеотидов левее сайта полиаденилирования. Делеции или мутации в гексамере AAUAAA препятствуют образованию полиаденилированного 3′-конца. И для разрезания цепи РНК, и для ее полиаденилирования необходим специфический сигнал.

Получение системы полиаденилирования *in vitro* облегчает изучение всех этих процессов. Структура и функции комплекса, осуществляющего 3′-процессинг, проиллюстрированы на РИС. 26.35. Формирование корректного 3′-конца требует участия **эндонуклеазы**, состоящей из двух субъединиц — CFI и CFII — и осуществляющей расщепление РНК. Синтез поли(A)-участка осуществляется **поли(A)-полимеразой (PAP).** *Компонент специфичности* (CPSF) распознает последовательность AAUAAA и направляет все остальные активности. Фактор стимулирования, CstF, связывается с G-U-богатым регионом, расположенным после сайта расщепления.

Фактор специфичности состоит из четырех субъединиц, связывающихся с последовательностью AAUAAA. Каждая индивидуальная субъединица представляет собой белок с РНК-связывающим мотивом, обладающий способностью неспецифично связывать-

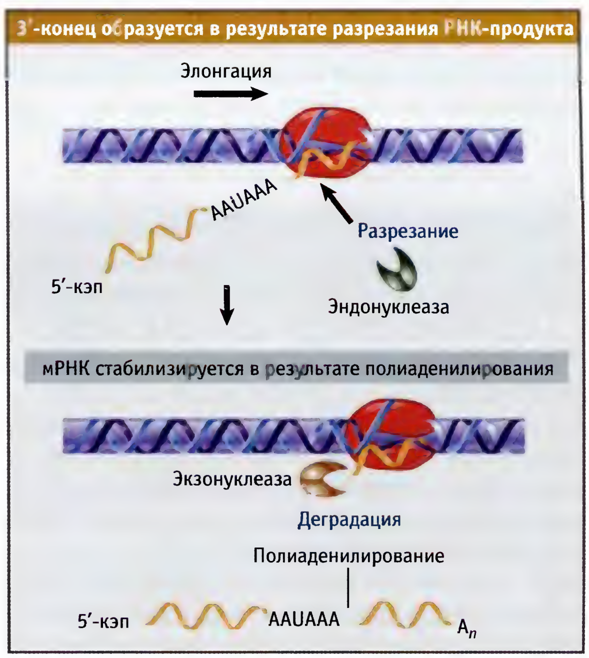


РИС. 26.34 Последовательность AAUAAA является сигналом для расщепления молекулы РНК. При этом образуется 3′-конец, который впоследствии полиаденилируется

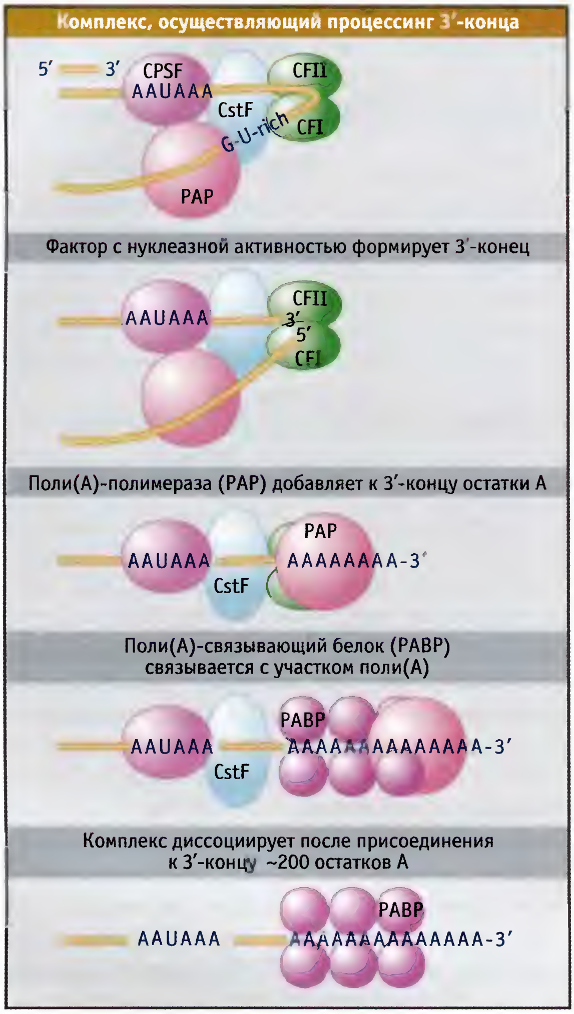


РИС. 26.35 **Комплекс, отвечающий за процессинг 3′-конца, обладает несколькими активностями. Белки CPSF и CstF состоят из нескольких субъединиц; другие компоненты являются мономерами. Общая масса комплекса составляет примерно 900 кДа**

ся с РНК. Специфический AAUAAA-связывающий сайт формируется в результате белок-белковых взаимодействий между субъединицами. CPSF прочно связывается с последовательностью AAUAAA только в том случае, если фактор CstF также присутствует в G–U-богатом участке.

Как реакция расщепления, так и реакция полиаденилирования требуют участия фактора специфичности. Он существует в комплексе с эндонуклеазой и поли(A)-полимеразой, поэтому процессы расщепления и полиаденилирования тесно связаны.

Двух компонентов комплекса — факторов CFI и CFII (факторы расщепления I и II) — вместе с фактором специфичности необходимо и достаточно для осуществления эндонуклеазной реакции.

Поли(A)-полимераза характеризуется неспецифической каталитической активностью. Однако в комплексе с другими компонентами она осуществляет реакцию, специфичную для РНК, содержащую в своем составе последовательность AAUAAA. Реакция полиаденилирования происходит в две стадии. Сначала к 3′-концу присоединяется короткая (порядка 10 оснований) последовательность олиго(A). Эта стадия зависит исключительно от AAUAAA-последовательности, работа РНК-полимеразы направляется фактором специфичности. Во время второй стадии концевая последовательность олиго(A) увеличивается до ~200 нуклеотидов. Эта реакция требует участия другого фактора стимулирования, который распознает олиго(A) и, координируя работу РНК-полимеразы, обеспечивает ее удлинение с 3′-конца.

Без участия дополнительных факторов механизм работы поли(A)-полимеразы демонстрирует разовый характер, фермент по одному добавляет к 3′-концу основания A, диссоциируя после каждого добавленного нуклеотида от синтезируемой цепи РНК. Однако в присутствии CPSF и PABP (поли(A)-связывающий белок) характер синтеза цепи олиго(A) сменяется на процессивный (непрерывный). PABP — белок с молекулярной массой 33 кДа, стехиометрически связывающийся с цепью олиго(A). Длина концевого фрагмента олиго(A) контролируется PABP и составляет, как правило, порядка 200 нуклеотидов. Такой предел может объясняться связыванием критического числа молекул PABP с поли(A). PABP связывается с фактором инициации трансляции eIF4G, формируя закрытую петлю, в которую заключены оба конца мРНК (см. рис. 8.20 в разд. 8.9 «У эукариот множество факторов инициации объединены в комплекс»).

Полиаденилирование в значительной мере определяет функции мРНК. Оно может влиять на стабильность молекулы и на процесс инициации трансляции (см. разд. 7.10 «3′-конец полиаденилирован»). У некоторых организмов во время эмбрионального развития последовательность поли(A) участвует в контроле трансляции. Полиаденилированная мРНК в этом случае активирует трансляцию, деаденилированная — ингибирует. Полиаденилирование мРНК в цитоплазме эмбрионов *Xenopus* зависит от специфического *цис*-активного элемента (CPE) на 3′-конце молекулы. Этим элементом является другая AU-последовательность — UUUUUAU.

В эмбрионах *Xenopus* обнаружено по крайней мере два *цис*-активных элемента, способных запустить деаденилирование. Эмбриональный элемент деаденилирования — это 17-нуклеотидная последовательность. ARE элементы представляют собой AU-богатые последовательности и, как правило, содержат тандемные повторы AUUUA. В деградацию мРНК вовлечена поли(A)-специфичная РНКаза (PARN). Не всегда деаденилирование запускается с помощью специфичных элементов; в некоторых случаях (включая процессы нормальной деградации мРНК в результате старения), последовательность поли(A) будет деградировать до тех пор, пока не будет специальным образом стабилизирована.

26.20 Расщепление и формирование 3′-конца мРНК гистонов происходят при участии малых РНК

Ключевые понятия:

- мРНК гистонов не полиаденилированы; их 3′-конец образуется в результате реакции расщепления, зависящей от структуры мРНК.
- Условием осуществления реакции расщепления является связывание SLBP с структурой типа стебель-петля на мРНК и спаривание U7 мяРНК с прилегающим к этой структуре одноцепочечным регионом.



РИС. 26.36 **Формирование 3′-конца мРНК гистона Н3 зависит от консервативной шпильки и последовательности, взаимодействующей с U7 мяРНК**

Некоторые мРНК не полиаденилированы. Механизм формирования их 3′-концов отличается от координированных реакций расщепления/полиаденилирования. Наиболее яркими представителями таких мРНК являются мРНК гистоновых белков, синтезирующиеся во время репликации ДНК. Процесс формирования 3′-концов зависит от вторичной структуры этих РНК. В 3′-концевой части молекулы находится высококонсервативная структура типа стебель-петля, стебель которой образован шестью, а петля — четырьмя нуклеотидами. Расщепление происходит четырьмя-пятью основаниями правее структуры типа стебель-петля. Для осуществления этой реакции требуется участие двух факторов: белка, связывающего структуру стебель-петля (SLBP), и U7 мяРНК, спаривающейся с пурин-богатым участком (histone downstream element или HDE), расположенной примерно на 10 нуклеотидов правее сайта расщепления.

Мутации, препятствующие формированию двухцепочечного стебля структуры стебель-петля, нарушают и процесс формирования 3′-конца РНК. Вторичные мутации, восстанавливающие способность к образованию дуплексной структуры (при этом не обязательно восстанавливающие исходную последовательность), возвращают и способность к образованию 3′-конца. Таким образом, *особая вторичная структура играет в процессе бóльшую роль, чем определенная нуклеотидная последовательность*. SLBP связывается со структурой стебель-петля, после чего взаимодействует с U7 мяРНП, упрочняя ее взаимодействие с правее расположенным сайтом связывания U7 мяРНК. U7 мяРНП –мяРНП небольшого размера, состоящая из 63-нуклеотидной мяРНК и нескольких белков (кроме Sm-белков; см. разд. 26.5 «мяРНК участвуют в сплайсинге»).

Реакция между мяРНК гистона Н3 и U7 мяРНК изображена на РИС. 26.36. Участки, которые соответствуют шпильке и HDE (взаимодействующему с U7 мяРНК), консервативны для Н3 мРНК нескольких видов. На 5′-конце U7 мяРНК расположена последовательность, связывающаяся с консервативной последовательностью гистоновой мРНК. Процессинг 3′-конца подавляется мутациями в HDE, препятствующими спариванию с U7 мяРНК. Компенсаторные мутации в U7 мяРНК, восстанавливающие комплементарность, восстанавливают и процессинг. Это свидетельствует о том, что U7 мяРНК осуществляет свою функцию, связываясь с гистоновой мРНК. Последовательность HDE различается у разных гистоновых мРНК, поэтому связывание с мяРНК само по себе не обладает достаточной стабильностью, но требует взаимодействия с SLBP.

Расщепление с образованием 3′-конца происходит на строго определенном расстоянии от сайта, узнаваемого U7 мяРНК. Это позволяет предположить, что мяРНК участвует в распознавании места расщепления. Фактор(ы), напрямую ответственные за расщепление, пока не установлены.

26.21 рРНК образуются в результате расщепления

Ключевые понятия:

- Большие и малые рРНК образуются в результате расщепления общего предшественника.

Все основные рРНК синтезируются как часть одного общего первичного транскрипта, в результате процессинга которого образуются зрелые продукты. Предшественник содержит в своем составе последовательности 18S, 5.8S и 28S рРНК. У высших эукариот предшественник, получивший свое название в соответствии с коэффициентом седиментации, называется **45S РНК**. У низших эукариот он меньше (у дрожжей — 35S).

Зрелые рРНК высвобождаются из состава предшественника в результате комбинации реакций расщепления и подравнивания. На приведен основной механизм, свойственный дрожжам. Порядок событий может различаться, однако основные реакции свойственны всем эукариотам. Большая часть 5′-концов является пря-

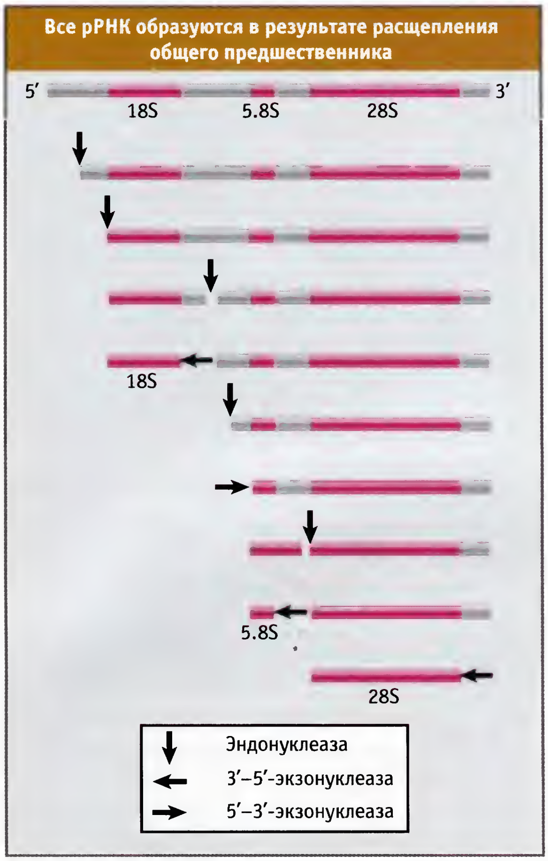


РИС. 26.37 Все зрелые эукариотические рРНК образуются в результате разрезания и подравнивания концов первичного транскрипта

мым результатом реакций расщепления. 3′-концы образуются в результате расщепления, за которым следует реакция 3′–5′-подравнивания.

В процессинге рРНК задействованы многие рибонуклеазы, в том числе и экзосома, представляющая собой комплекс нескольких экзонуклеаз, также обеспечивающих деградацию мРНК (см. разд. 7.13 «В деградацию мРНК вовлечены ферменты с различными активностями»). Мутации в отдельных ферментах обычно не препятствуют процессингу. Это означает, что их активность избыточна и для образования зрелых молекул рРНК могут использоваться различные комбинации реакций расщепления.

Транскрипционная единица рРНК всегда существует в виде большого количества геномных копий. Эти копии организованы в тандемные повторы (см. разд. 6.9 «Множественные гены рРНК сохраняют постоянство своей последовательности»).

5S рРНК транскрибируется с отдельного гена РНК-полимеразой III. Гены 5S рРНК организованы в кластеры, однако отделены от генов основных рРНК. (У дрожжей 5S ген ассоциирован с каждой транскрипционной единицей основных рРНК, однако транскрибируется независимо.)

В этом заключается разница между организацией предшественника у эукариот и у бактерий. Последовательность, соответствующая 5.8S рРНК, формирует 5′-конец большой (23S) рРНК, процессинга между этими последовательностями не происходит. Кроме того, как показано на РИС. 26.38, предшественник содержит последовательность 5S РНК и одной или двух тРНК. У *E. coli* семь *rrn* оперонов распределены по всему геному; четыре *rrn* локуса между последовательностями 16S и 23S РНК содержат ген одной тРНК, три других содержат в этом регионе два гена тРНК. Дополнительные гены тРНК могут располагаться между последовательностью 5S РНК и 3′-концом молекулы. Таким образом, ход реакций процессинга, обеспечивающих образование зрелых продуктов, зависит от состава конкретного *rrn* локуса.

При процессинге как про-, так и эукариотических рРНК рибосомальные белки (возможно, и другие белки) связываются с предшественником таким образом, что субстратом для процессинга становится не свободная РНК, а рибонуклеопротеиновый комплекс.

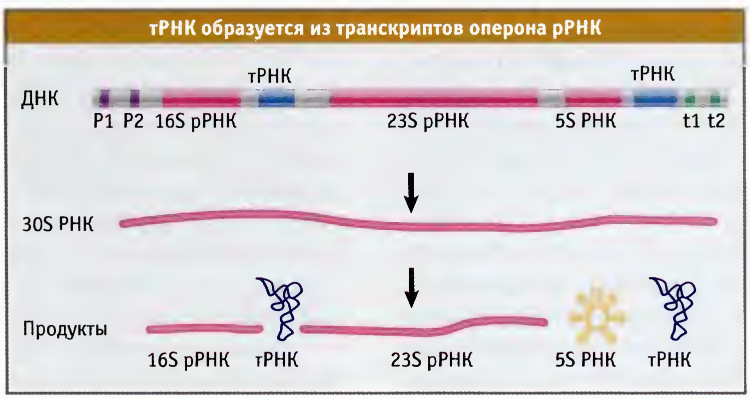


РИС. 26.38 *rrn* оперон *E. coli* содержит гены рРНК и тРНК. Точная длина транскрипта зависит от того, какой именно промотор (P) и терминатор (t) задействованы. Каждый РНК-продукт высвобождается из первичного транскрипта в результате разрезания в определенных сайтах

26.22 Малые РНК участвуют в процессинге рРНК

Ключевые понятия:

- C/D-группа snoРНК необходима для метилирования рибозы в 2′-положении.
- H/ACA-группа snoРНК необходима для превращения уридина в псевдоуридин.
- В каждом случае snoРНК спаривается с последовательностью рРНК, содержащей основания-мишени. При этом формируется структура, являющаяся субстратом для модификации.

Процессинг и модификация рРНК требуют участия малых РНК, названных snoРНК (малые ядрышковые РНК). В геноме дрожжей (*S. cerevisiae*) обнаружены 71 snoРНК. Они ассоциированы с белком фибрилларином, являющимся одним из основных компонентов ядрышка (регион ядра, в котором происходит транскрипция генов рРНК). Некоторые snoРНК участвуют в расщеплении предшественника рРНК; примером может служить U3 snoРНК, необходимая для первой реакции расщепления предшественника рРНК у дрожжей и *Xenopus*. Мы не знаем, какую роль snoРНК играет в расщеплении. Возможно, она обеспечивает формирование особой вторичной структуры рРНК, узнаваемой эндонуклеазой.

Две группы snoРНК обеспечивают модификацию оснований рРНК. Члены каждой из этих групп определяются по короткой консервативной последовательности и общим чертам вторичной структуры.

C/D-группа snoРНК необходима для метилирования 2′-О-атома рибозы. В консервативных положениях рРНК позвоночных встречается более 100 2′-О-метильных групп. C/D-группа получила свое название благодаря двум коротким консервативным мотивам, названным сайтами C и D. Каждая snoРНК рядом с D сайтом содержит последовательность, комплементарную метилируемым регионам 18S или 28S рРНК. Утрата какой-либо snoРНК ведет к нарушению метилирования в том участке рРНК, которому она была комплементарна.

На рис. 26.39 показано, как snoРНК спаривается с рРНК, формируя дуплексный регион, распознаваемый как субстрат для метилирования. Метилирование происходит внутри области комплементарности, возможное положение нуклеотида-мишени ограничено пятью основаниями на 5′-конце D сайта. По всей видимости, специфичность реакции метилирования определяется каждой конкретной snoРНК; к настоящему моменту охарактеризовано порядка 40 snoРНК. Метилаза(ы) пока не охарактеризованы; одна из возможностей состоит в том, что частично метилазная активность обеспечивается самой snoРНК.

Другая группа snoРНК участвует в синтезе псевдоуридина. В дрожжевых рРНК обнаружно 43 ψ-остатка, в рРНК позвоночных — порядка 100. Синтез псевдоуридина обеспечивается реакцией, представленной на рис. 26.40. Связь между атомом N1 уридиловой кислоты и рибозой разрушается, основание разворачивается, и формируется связь между атомом C5 и сахаром.

Образование псевдоуридина в рРНК происходит при участии snoРНК группы H/ACA, наличие которой характерно для примерно 20 snoРНК. Свое название группа получила благодаря триплету ACA, расположенному в трех нуклеотидах от 3′-конца молекулы, и частично консервативной последовательности (H-сайту), находящейся между двумя структурами типа стебель-петля. Все 20 snoРНК внутри стебля каждой из этих структур имеют последовательность, комплементарную рРНК.

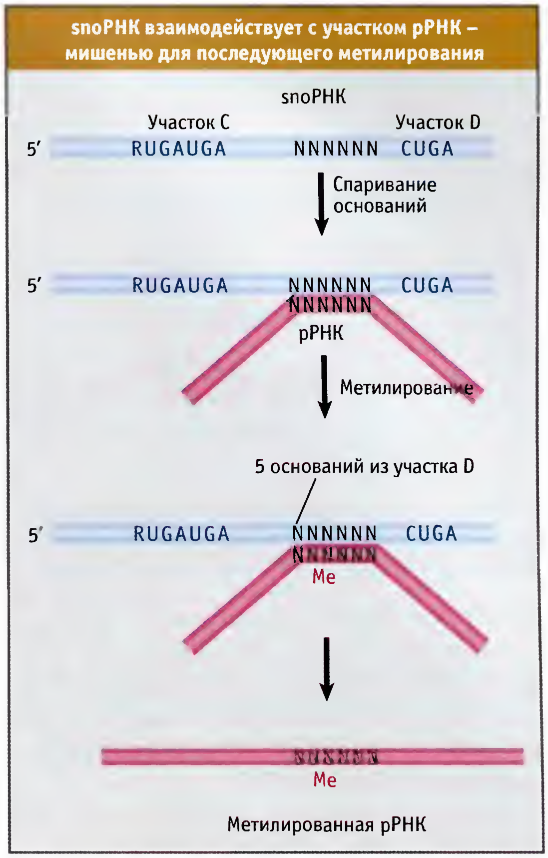


Рис. 26.39 snoРНК взаимодействует с тем регионом мРНК, который впоследствии метилируется

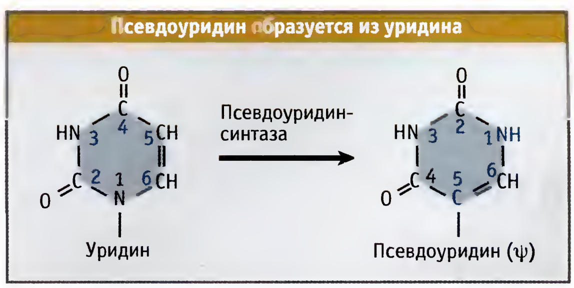


Рис. 26.40 Псевдоуридин образуется из уридина в результате замещения связи N1-сахар на связь C5-сахар и поворота основания относительно сахарного остатка

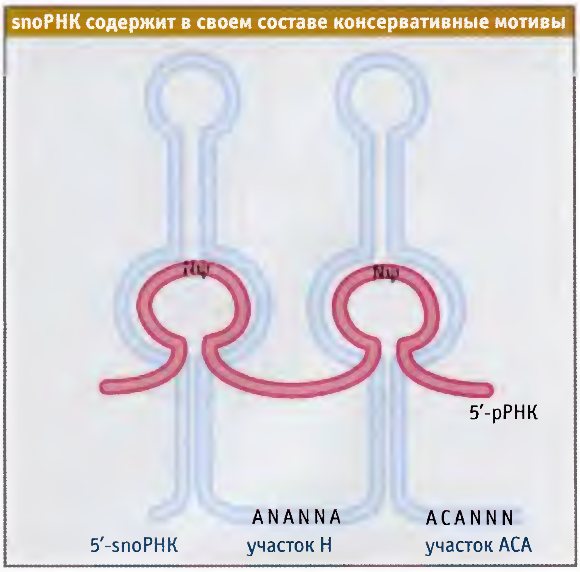


РИС. 26.41 H/ACA snoPHK содержит в своем составе две короткие консервативные последовательности и две шпилечные структуры. В стебле каждой из этих структур имеются регионы, комплементарные рРНК. Псевдоуридин образуется в результате преобразований неспаренного уридина, расположенного в комплементарной snoPHK области рРНК

На РИС. 26.41 изображена структура, образующаяся при спаривании snoPHK и рРНК. Внутри каждого региона комплементарности есть два неспаренных основания, одним из которых является уридин, превратившийся в псевдоуридин.

H/ACA snoPHK ассоциированы с ядрышковым белком Gar1p, необходимым для формирования псевдоуридина. Точная функция этого белка не известна. Описанные псевдоуридин-синтазы являются белками, функционирующими без участия РНК-кофактора. Синтаза, участвующая в snoPHK-опосредованном синтезе псевдоуридина, не определена.

Участие U7 мяРНК в формировании 3'-конца молекул рРНК, а также роль snoPHK в процессинге и модификации рРНК согласуются с данными, которые мы приводим в гл. 27 «Каталитическая РНК». Многие, а возможно, и все события процессинга РНК зависят от РНК-РНКовых взаимодействий. Общим свойством всех РНК, входящих в состав частиц, является способность спариваться с короткой последовательностью в составе субстратной РНК.

26.23 Резюме

Сплайсинг представляет собой удаление интронов и соединение экзонов в молекулу зрелой мРНК. По условиям, которые требуются для протекания реакции *in vitro,* а также образующимся в ходе реакции промежуточным продуктам, все реакции сплайсинга можно разделить по крайней мере на четыре типа: сплайсинг ядерных интронов, интронов группы I, интронов группы II, интронов тРНК. Каждая реакция представляет собой изменение организации молекулы РНК и поэтому является *цис*-активным событием.

Сплайсинг пре-мРНК происходит по предпочтительному, однако не обязательному механизму. Осуществление реакции требует участия лишь очень коротких консервативных последовательностей; остальная часть последовательности интрона в обеспечении сплайсинга участия не принимает. Все 5'- и 3'-сайты сплайсинга, по-видимому, функционально эквивалентны. Участвующие в сплайсинге последовательности ограничивают концы интрона и описываются правилом GU–AG. Сайт разветвления, представленный последовательностью UACUAAC у дрожжей и менее консерватиной последовательностью у млекопитающих, также принимает участие в осуществлении реакции. 5'-сайт сплайсинга задействован в формировании структуры типа лассо, с помощью которого через 5'–2'-связь соединяются GU-конец интрона и основание А в положении 6 сайта разветвления. После этого 3'-OH конец экзона атакует 3'-сайт сплайсинга, при этом происходит сшивание экзонов и высвобождение интрона в виде лассо. Обе реакции по химической природе являются реакциями трансэтерификации, в которых число связей остается постоянным. Некоторые стадии реакции происходят при участии АТФ, требуемой, возможно, для обеспечения конформационных изменений РНК и/или белковых компонентов. Формирование лассо обеспечивает узнавание 3'-сайта сплайсинга. При альтернативном сплайсинге белковые факторы либо стимулируют использование новых вариантов сшивки экзонов, либо блокируют использование обычных вариантов.

Сплайсинг пре-мРНК происходит при участии сплайсосомы — большой частицы, собирающей в своем составе консенсусные последовательности в активную конформацию. Чаще всего сплайсосома формируется в ходе процесса, получившего название распознавание интронов, заключающегося в распознавании 5'-сайта сплайсинга, сайта разветвления 3'-сайта сплайсинга. Альтернативным вариантом является распознавание экзонов, при этом происходит распознавание 5'-сайтов субстратного интрона и следующего за ним интрона. Сборка сплайсосомы проходит несколько стадий от комлекса E (комплекса, определяющего вырезание данного интрона), содержащего в своем составе U1 мяРНП и факторы сплайсинга, к A и B комплексам, образующимся при присоединении дополнительных компонентов.

Сплайсосома состоит из U1, U2, U4/U6 и U5 мяРНП и дополнительных факторов сплайсинга. В составе U1, U2 и U5 выделяют единственную мяРНК и несколько белков; U4/U6 мяРНП представлена двумя мяРНК и белками. Ряд белков встречается в составе всех мяРНП. мяРНП распознают консенсусные последовательности. U1 мяРНП спаривается с 5'-сайтом сплайсинга, U2 мяРНП — с сайтом разветвления, а U5 мяРНП действует на 5'-сайт. После освобождения из комплекса U4/U6 U6 мяРНК взаимодействует с U2; в результате этого взаимодействия формиру-

ется каталитический центр сплайсинга. Альтернативный набор мяРНП обеспечивает выполнение аналогичных функций при сплайсинге U12-зависимых интронов. Молекулы мяРНК, возможно, играют каталитическую роль в сплайсинге и других реакциях процессинга.

В ядрышке две группы snoPHK спариваются с теми сайтами pPHK, которые в дальнейшем будут подвергаться модификации. Мишенью для связывания со snoPHK группы C/D являются сайты метилирования, а для связывания со snoPHK группы ACA — сайты превращения уридина в псевдоуридин.

Сплайсинг является, как правило, внутримолекулярной реакцией. Однако некоторым организмам (трипаносомы, нематоды) свойственен *транс*-сплайсинг (межмолекулярный сплайсинг). *Транс*-сплайсинг представляет собой взаимодействие малой СЛ РНК и пре-мРНК. СЛ РНК напоминает U1 мяРНК и совмещает в себе ее функции и функцию предоставления экзона для сплайсинга. У червей обнаружено два типа СЛ РНК: один используется в сплайсинге с участием 5′-конца мРНК, другой — с участием внутренних сайтов.

Для интронов группы II, как и для ядерных интронов, характерно образование лассо в качестве промежуточного продукта. Их отличительной особенностью является способность к автокатализу. Сплайсинг этих интронов следует правилу GT–AG, однако при этом формируется специфическая вторичная структура, удерживающая реакционные сайты в соответствующем положении.

Сплайсинг дрожжевой тРНК происходит в результате отдельных эндонуклеазной и лигазной реакций. Эндонуклеаза распознает вторичную (или третичную) структуру предшественника и разрезает его молекулу по обоим концам интрона. Две половины молекулы тРНК, образовавшиеся в результате вырезания интрона, сшиваются в присутствии АТФ.

Способность РНК-полимеразы II терминировать свою работу не описана, 3′-концы ее транскриптов образуются в результате расщепления синтезируемой цепи РНК. Последовательность AAUAAA, расположенная на 11–30 оснований левее сайта расщепления, является сигналом для осуществления реакций расщепления и полиаденилирования. Эндонуклеаза и поли(А)-полимераза ассоциированы в один комплекс с факторами, обеспечивающими подтверждение специфичности сигнала AAUAAA. Транскрипция терминируется после события расщепления, когда экзонуклеаза, связанная с 5′-концом синтезируемой цепи РНК, догоняет РНК-полимеразу.

Литература

26.1 Введение

Обзор

Dreyfuss, G., Kim, V. N., and Kataoka, N. (2002). Messenger-PHK-binding proteins and the messages they carry. *Nat. Rev. Mol. Cell Biol.* 3, 195–205.
Dreyfuss, G. et al. (1993). hnRNP proteins and the biogenesis of mPHK. *Annu. Rev. Biochem.* 62, 289–321.
Lewin, B. (1975). Units of transcription and translation: sequence components of hnPHK and mPHK. *Cell* 4, 77–93.

26.2 Сайты сплайсинга ядерных генов представляют собой короткие последовательности

Обзор

Padgett, R. A. (1986). Splicing of messenger PHK precursors. *Annu. Rev. Biochem.* 55, 1119–1150.
Sharp, P. A. (1987). Splicing of mPHK precursors. *Science* 235, 766–771.

26.4 В ходе сплайсинга мРНК образуется структура типа лассо

Обзор

Sharp, P. A. (1994). Split genes and PHK splicing. *Cell* 77, 805–815.
Weiner, A. (1993). mPHK splicing and autocatalytic introns: distant cousins or the products of chemical determinism. *Cell* 72, 161–164.

Исследование

Reed, R. and Maniatis, T. (1985). Intron sequences involved in lariat formation during pre-mPHK splicing. *Cell* 41, 95–105.
Reed, R. and Maniatis, T. (1986). A role for exon sequences and splice-site proximity in splice-site selection. *Cell* 46, 681–690.
Zhuang, Y. A., Goldstein, A. M., and Weiner, A. M. (1989). UACUAAC is the preferred branch site for mammalian mPHK splicing. *Proc. Natl. Acad. Sci. USA* 86, 2752–2756.

26.5 мяРНК участвуют в сплайсинге

Обзор

Guthrie, C. (1991). Messenger PHK splicing in yeast: clues to why the spliceosome is a ribonucleoprotein. *Science* 253, 157–163.
Guthrie, C. and Patterson, B. (1988). Spliceosomal snPHKs. *Annu. Rev. Genet.* 22, 387–419.
Maniatis, T. and Reed, R. (1987). The role of small nuclear ribonucleoprotein particles in pre-mPHK splicing. *Nature* 325, 673–678.

Исследование

Grabowski, P. J., Seiler, S. R., and Sharp, P. A. (1985). A multicomponent complex is involved in the splicing of messenger PHK precursors. *Cell* 42, 345–353.
Zhou, Z., Licklider, L. J., Gygi, S. P., and Reed, R. (2002). Comprehensive proteomic analysis of the human spliceosome. *Nature* 419, 182–185.

26.6 U1 мяРНП инициирует сплайсинг

Обзор

Brow, D. A. (2002). Allosteric cascade of spliceosome activation. *Annu. Rev. Genet.* 36, 333–360.

Исследование

Abovich, N. and Rosbash, M. (1997). Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. *Cell* 89, 403–412.
Berglund, J. A., Chua, K., Abovich, N., Reed, R., and Rosbash, M. (1997). The splicing factor BBP interacts specifically with the pre-mPHK branchpoint sequence UACUAAC. *Cell* 89, 781–787.

Burgess, S., Couto, J. R., and Guthrie, C. (1990). A putative ATP binding protein influences the fidelity of branchpoint recognition in yeast splicing. *Cell* 60, 705–717.

Parker, R., Siliciano, P. G., and Guthrie, C. (1987). Recognition of the TACTAAC box during mPHK splicing in yeast involves base pairing to the U2-like snPHK. *Cell* 49, 229–239.

Singh, R., Valcárcel, J., and Green, M. R. (1995). Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. *Science* 268, 1173–1176.

Wu, S., Romfo, C. M., Nilsen, T. W., and Green, M. R. (1999). Functional recognition of the 3′ splice site AG by the splicing factor U2AF35. *Nature* 402, 832–835.

Zamore, P. D. and Green, M. R. (1989). Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. *Proc. Natl. Acad. Sci. USA* 86, 9243–9247.

Zhang, D. and Rosbash, M. (1999). Identification of eight proteins that cross-link to pre-mPHK in the yeast commitment complex. *Genes Dev.* 13, 581–592.

Zhuang, Y. and Weiner, A. M. (1986). A compensatory base change in U1 snPHK suppresses a 5 splice site mutation. *Cell* 46, 827–835.

26.7 Формирование комплекса E может происходить в результате распознавания сплайсинга как интронов, так и экзонов

Исследование

Bruzik, J. P. and Maniatis, T. (1995). Enhancer-dependent interaction between 5′ and 3′ splice sites in *trans*. *Proc. Natl. Acad. Sci. USA* 92, 7056–7059.

26.8 Пять мяРНП формируют сплайсосому

Обзор

Krämer, A. (1996). The structure and function of proteins involved in mammalian pre-mPHK splicing. *Annu. Rev. Biochem.* 65, 367–409.

Madhani, H. D. and Guthrie, C. (1994). Dynamic PHK-PHK interactions in the spliceosome. *Annu. Rev. Genet.* 28, 1–26.

Исследование

Lamond, A. I. (1988). Spliceosome assembly involves the binding and release of U4 small nuclear ribonucleoprotein. *Proc. Natl. Acad. Sci. USA* 85, 411–415.

Lesser, C. F. and Guthrie, C. (1993). Mutations in U6 snPHK that alter splice site specificity: implications for the active site. *Science* 262, 1982–1988.

Madhani, H. D. and Guthrie, C. (1992). A novel base-pairing interaction between U2 and U6 snPHKs suggests a mechanism for the catalytic activation of the spliceosome. *Cell* 71, 803–817.

Newman, A. and Norman, C. (1991). Mutations in yeast U5 snPHK alter the specificity of 5′ splice site cleavage. *Cell* 65, 115–123.

Sontheimer, E. J. and Steitz, J. A. (1993). The U5 and U6 small nuclear PHKs as active site components of the spliceosome. *Science* 262, 1989–1996.

26.9 Аппарат альтернативного сплайсинга использует различные мяРНП

Исследование

Burge, C. B., Padgett, R. A., and Sharp, P. A. (1998). Evolutionary fates and origins of U12-type introns. *Mol. Cell* 2, 773–785.

Dietrich, R. C., Incorvaia, R., and Padgett, R. A. (1997). Terminal intron dinucleotide sequences do not distinguish between U2- and U12-dependent introns. *Mol. Cell* 1, 151–160.

Tarn, W.-Y. and Steitz, J. (1996). A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class AT-AC intron *in vitro*. *Cell* 84, 801–811.

26.10 Сплайсинг сопряжен с экспортом мРНК

Обзор

Dreyfuss, G., Kim, V. N., and Kataoka, N. (2002). Messenger-PHK-binding proteins and the messages they carry. *Nat. Rev. Mol. Cell Biol.* 3, 195–205.

Reed, R. and Hurt, E. (2002). A conserved mPHK export machinery coupled to pre-mPHK splicing. *Cell* 108, 523–531.

Исследование

Kataoka, N., Yong, J., Kim, V. N., Velazquez, F., Perkinson, R. A., Wang, F., and Dreyfuss, G. (2000). Pre-mPHK splicing imprints mPHK in the nucleus with a novel PHK-binding protein that persists in the cytoplasm. *Mol. Cell* 6, 673–682.

Le Hir, H., Gatfield, D., Izaurralde, E., and Moore, M. J. (2001). The exon-exon junction complex provides a binding platform for factors involved in mPHK export and nonsense-mediated mPHK decay. *EMBO J.* 20, 4987–4997.

Le Hir, H., Izaurralde, E., Maquat, L. E., and Moore, M. J. (2000). The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mPHK exon-exon junctions. *EMBO J.* 19, 6860–6869.

Luo, M. J. and Reed, R. (1999). Splicing is required for rapid and efficient mPHK export in metazoans. *Proc. Natl. Acad. Sci. USA* 96, 14937–14942.

Luo, M. L., Zhou, Z., Magni, K., Christoforides, C., Rappsilber, J., Mann, M., and Reed, R. (2001). Pre-mPHK splicing and mPHK export linked by direct interactions between UAP56 and Aly. *Nature* 413, 644–647.

Reichert, V. L., Le Hir, H., Jurica, M. S., and Moore, M. J. (2002). 5′ exon interactions within the human spliceosome establish a framework for exon junction complex structure and assembly. *Genes Dev.* 16, 2778–2791.

Rodrigues, J. P., Rode, M., Gatfield, D., Blencowe, B., Blencowe, M., and Izaurralde, E. (2001). REF proteins mediate the export of spliced and unspliced mPHKs from the nucleus. *Proc. Natl. Acad. Sci. USA* 98, 1030–1035.

Strasser, K. and Hurt, E. (2001). Splicing factor Sub2p is required for nuclear mPHK export through its interaction with Yra1p. *Nature* 413, 648–652.

Zhou, Z., Luo, M. J., Straesser, K., Katahira, J., Hurt, E., and Reed, R. (2000). The protein Aly links pre-messenger-PHK splicing to nuclear export in metazoans. *Nature* 407, 401–405.

26.11 Автосплайсинг интронов группы II происходит через формирование лассо

Обзор

Michel, F. and Ferat, J.-L. (1995). Structure and activities of group II introns. *Annu. Rev. Biochem.* 64, 435–461.

26.12 Альтернативный сплайсинг — это реализация различных вариантов сшивания экзонов между собой

Обзор

Green, M. R. (1991). Biochemical mechanisms of constitutive and regulated pre-mPHK splicing. *Annu. Rev. Cell Biol.* 7, 559–599.

Исследование

Handa, N., Nureki, O., Kurimoto, K., Kim, I., Sakamoto, H., Shimura, Y., Muto, Y., and Yokoyama, S. (1999). Structural basis for recognition of the tra mPHK precursor by the Sex-lethal protein. *Nature* 398, 579–585.

Lynch, K. W. and Maniatis, T. (1996). Assembly of specific SR protein complexes on distinct regulatory elements of the *Drosophila* doublesex splicing enhancer. *Genes Dev.* 10, 2089–2101.

Sun, Q., Mayeda, A., Hampson, R. K., Krainer, A. R., and Rottman, F. M. (1993). General splicing factor SF2/ASF promotes altePHKtive splicing by binding to an exonic splicing enhancer. *Genes Dev.* 7, 2598–2608.

Tian, M. and Maniatis, T. (1993). A splicing enhancer complex controls altePHKtive splicing of doublesex pre-mPHK. *Cell* 74, 105–114.

Wu, J. Y. and Maniatis, T. (1993). Specific interactions between proteins implicated in splice site selection and regulated altePHKtive splicing. *Cell* 75, 1061–1070.

26.13 В реакциях *транс*-сплайсинга участвуют малые РНК

Обзор

Nilsen, T. (1993). Trans-splicing of nematode pre-mPHK. *Annu. Rev. Immunol.* 47, 413–440.

Исследование

Blumenthal, T., Evans, D., Link, C. D., Guffanti, A., Lawson, D., Thierry-Mieg, J., Thierry-Mieg, D., Chiu, W. L., Duke, K., Kiraly, M., and Kim, S. K. (2002). A global analysis of *C. elegans* operons. *Nature* 417, 851–854.

Hannon, G. J. et al. (1990). *trans*-splicing of nematode pre-mPHK *in vitro*. *Cell* 61, 1247–1255.

Huang, X. Y. and Hirsh, D. (1989). A second *trans*-spliced PHK leader sequence in the nematode *C. elegans*. *Proc. Natl. Acad. Sci. USA* 86, 8640–8644.

Krause, M. and Hirsh, D. (1987). A *trans*-spliced leader sequence on actin mPHK in *C. elegans*. *Cell* 49, 753–761.

Murphy, W. J., Watkins, K. P., and Agabian, N. (1986). Identification of a novel Y branch structure as an intermediate in trypanosome mPHK processing: evidence for *trans*-splicing. *Cell* 47, 517–525.

Sutton. R. and Boothroyd, J. C. (1986). Evidence for *trans*-splicing in trypanosomes. *Cell* 47, 527–535.

26.15 Эндонуклеаза сплайсинга распознает тРНК

Исследование

Baldi, I. M. et al. (1992). Participation of the intron in the reaction catalyzed by the *Xenopus* tPHK splicing endonuclease. *Science* 255, 1404–1408.

Diener, J. L. and Moore, P. B. (1998). Solution structure of a substrate for the archaeal pre-tPHK splicing endonucleases: the bulge-helix-bulge motif. *Mol. Cell* 1, 883–894.

Di Nicola Negri, E., Fabbri, S., Bufardeci, E., Baldi, M. I., Mattoccia, E., and Tocchini-Valentini, G. P. (1997). The eucaryal tPHK splicing endonuclease recognizes a tripartite set of PHK elements. *Cell* 89, 859–866.

Kleman-Leyer, K., Armbruster, D. W., and Daniels, C. J. (2000). Properties of H. volcanii tPHK intron endonuclease reveal a relationship between the archaeal and eucaryal tPHK intron processing systems. *Cell* 89, 839–847.

Lykke-Andersen, J. and Garrett, R. A. (1997). PHK-protein interactions of an archaeal homotetrameric splicing endoribonuclease with an exceptional evolutionary history. *EMBO J.* 16, 6290–6300.

Mattoccia, E. et al. (1988). Site selection by the tPHK splicing endonuclease of *X. laevis*. *Cell* 55, 731–738.

Reyes, V. M. and Abelson, J. (1988). Substrate recognition and splice site determination in yeast tPHK splicing. *Cell* 55, 719–730.

Trotta, C. R., Miao, F., Arn, E. A., Stevens, S. W., Ho, C. K., Rauhut, R., and Abelson, J. N. (1997). The yeast tPHK splicing endonuclease: a tetrameric enzyme with two active site subunits homologous to the archaeal tPHK endonucleases. *Cell* 89, 849–858.

26.17 Каскад реакций, запускаемых в ответ на накопление белков, связан с процессом сплайсинга тРНК

Исследование

Gonzalez, T. N., Sidrauski, C., Dorfler, S., and Walter, P. (1999). Mechanism of non-spliceosomal mPHK splicing in the unfolded protein response pathway. *EMBO J.* 18, 3119–3132.

Sidrauski, C., Cox, J. S., and Walter, P. (1996). tPHK ligase is required for regulated mPHK splicing in the unfolded protein response. *Cell* 87, 405–413.

Sidrauski, C. and Walter, P. (1997). The transmembrane kinase Ire1p is a site-specific endonuclease that initiates mPHK splicing in the unfolded protein response. *Cell* 90, 1031–1039.

26.19 3′-концы мРНК образуются в результате расщепления и полиаденилирования первичного транскрипта

Обзор

Wahle, E. and Keller, W. (1992). The biochemistry of 3′-end cleavage and polyadenylation of messenger PHK precursors. *Annu. Rev. Biochem.* 61, 419–440.

Исследование

Bouvet, P., Omilli, F., Arlot-Bonnemains, Y., Legagneux, V., Roghi, C., Bassez, T., and Osborne, H. B. (1994). The deadenylation conferred by the 3′ untranslated region of a developmentally controlled mPHK in Xenopus embryos is switched to polyadenylation by deletion of a short sequence element. *Mol. Cell Biol.* 14, 1893–1900.

Conway, L. and Wickens, M. (1985). A sequence downstream of AAUAAA is required for formation of SV40 late mPHK 3′ termini in frog oocytes. *Proc. Natl. Acad. Sci. USA* 82, 3949–3953.

Fox, C. A., Sheets, M. D., and Wickens, M. P. (1989). Poly(A) addition during maturation of frog oocytes: distinct nuclear and cytoplasmic activities and regulation by the sequence UUUUUAU. *Genes Dev.* 3, 2151–2162.

Gil, A. and Proudfoot, N. (1987). Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit β-globin mPHK 3′ end formation. *Cell* 49, 399–406.

Karner, C. G., Wormington, M., Muckenthaler, M., Schneider, S., Dehlin, E., and Wahle, E. (1998). The deadenylating nuclease (DAN) is involved in poly(A) tail removal during the meiotic maturation of *Xenopus* oocytes. *EMBO J.* 17, 5427–5437.

Kim, M., Krogan, N. J., Vasiljeva, L., Rando, O. J., Nedea, E., Greenblatt, J. F., and Buratowski, S. (2004). The yeast Rat1 exonuclease promotes transcription termination by РНК polymerase II. *Nature* 432, 517–522.

McGrew, L. L., Dworkin-Rastl, E., Dworkin, M. B., and Richter, J. D. (1989). Poly(A) elongation during *Xenopus* oocyte maturation is required for translational recruitment and is mediated by a short sequence element. *Genes Dev.* 3, 803–815.

Takagaki. Y., Ryner, L. C., and Manley, J. L. (1988). Separation and characterization of a poly(A) polymerase and a cleavage/specificity factor required for pre-mРНК polyadenylation. *Cell* 52, 731–742.

Voeltz, G. K. and Steitz, J. A. (1998). AUUUA sequences direct mРНК deadenylation uncoupled from decay during *Xenopus* early development. *Mol. Cell Biol.* 18, 7537–7545.

26.20 Расщепление и формирование 3′-конца мРНК гистонов происходит при участии малых РНК

Обзор

Birnstiel, M. L. (1985). Transcription termination and 3′ processing: the end is in site. *Cell* 41, 349–359.

Исследование

Bond, U. M., Yario, T. A., and Steitz, J. A. (1991). Multiple processing-defective mutations in a mammalian histone pre-mРНК are suppressed by compensatory changes in U7 РНК both *in vitro* and *in vivo*. *Genes Dev.* 5, 1709–1722.

Dominski, Z., Erkmann, J. A., Greenland, J. A., and Marzluff, W. F. (2001). Mutations in the РНК binding domain of stem-loop binding protein define separable requirements for РНК binding and for histone pre-mРНК processing. *Mol. Cell. Biol.* 21, 2008–2017.

Galli, G. et al. (1983). Biochemical complementation with РНК in the *Xenopus* oocyte: a small РНК is required for the generation of 3′ histone mРНК termini. *Cell* 34, 823–828.

Mowry, K. L. and Steitz, J. A. (1987). Identification of the human U7 snRNP as one of several factors involved in the 3′ end maturation of histone premessenger РНК's. *Science* 238, 1682–1687.

Wang, Z. F., Whitfield, M. L., Ingledue, T. C., Dominski, Z., and Marzluff, W. F. (1996). The protein that binds the 3′ end of histone mРНК: a novel РНК-binding protein required for histone pre-mРНК processing. *Genes Dev.* 10, 3028–3040.

26.21 рРНК образуются в результате расщепления

Обзор

Venema, J. and Tollervey, D. (1999). Ribosome synthesis in *S. cerevisiae*. *Annu. Rev. Genet.* 33, 261–311.

26.22 Малые РНК участвуют в процессинге рРНК

Исследование

Balakin, A. G., Smith, L., and Fournier, M. J. (1996). The РНК world of the nucleolus: two major families of small РНКs defined by different box elements with related functions. *Cell* 86, 823–834.

Bousquet-Antonelli, C., Henry, Y., G'elugne, J. P., Caizergues-Ferrer, M., and Kiss, T. (1997). A small nucleolar RNP protein is required for pseudouridylation of eukaryotic ribosomal РНКs. *EMBO J.* 16, 4770–4776.

Ganot, P., Bortolin, M. L., and Kiss, T. (1997). Site-specific pseudouridine formation in preribosomal РНК is guided by small nucleolar РНКs. *Cell* 89, 799–809.

Ganot, P., Caizergues-Ferrer, M., and Kiss, T. (1997). The family of box ACA small nucleolar РНКs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for РНК accumulation. *Genes Dev.* 11, 941–956.

Kass, S. et al. (1990). The U3 small nucleolar ribonucleoprotein functions in the first step of preribosomal РНК processing. *Cell* 60, 897–908.

Kiss-Laszlo, Z., Henry, Y., and Kiss, T. (1998). Sequence and structural elements of methylation guide snoРНКs essential for site-specific ribose methylation of pre-rРНК. *EMBO J.* 17, 797–807.

Kiss-Laszlo, Z. et al. (1996). Site-specific ribose methylation of preribosomal РНК: a novel function for small nucleolar РНКs. *Cell* 85, 1077–1068.

Ni, J., Tien, A. L., and Fournier, M. J. (1997). Small nucleolar РНКs direct site-specific synthesis of pseudouridine in rРНК. *Cell* 89, 565–573.

Каталитическая РНК

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

27.10 Редактирование РНК касается конкретных азотистых оснований

- В незрелых мРНК аполипопротеина В и рецепторов глутамата происходит сайт-специфичное дезаминирование; катализируемое цитидин- и аденозин-дезаминазами; реакции приводят к изменению смысла кодирующей области.

27.11 Редактирование транскриптов направляется специальными малыми молекулами РНК

- В митохондриях трипаносом имеет место экстенсивное редактирование РНК: транскрипты конкретных генов претерпевают многочисленные вставки и делеции уридинов.
- Две молекулы, транскрипт и наводящая РНК, взаимодействуют своими комплементарными участками; эти участки располагаются по обе стороны от последовательности-мишени редактирования.
- Матрицей для добавления (реже, удаления) уридинов служит наводящая РНК.
- Катализ редактирования осуществляется белковым комплексом, который обладает активностями эндонуклеазы, терминальной уридил-трансферазы и РНК-лигазы.

27.12 Сплайсинг белка — автокаталитическая реакция

- Участок полипептида, обозначаемый как интеин, может катализировать свой собственный выход из состава белка; при этом фланкирующие участки, экстеины, смыкаются друг с другом; это событие известно как сплайсинг белка.
- Катализатором сплайсинга белка служит интеин.
- У большинства интеинов есть две независимых активности: кроме участия в сплайсинге белка многие интеины проявляют себя как хоминг-эндонуклеазы.

27.13 Резюме

27.1 Введение

Согласно традиционным представлениям, любая энзиматическая активность (то есть способность к катализу биохимических реакций) всегда приписывалась белкам. Долгое время считалось, что только белки, в многообразии своей пространственной структуры и в силу своей разносторонней химической природы обладают особенной пластичностью и служат материалом, из которого эволюция «создает» активные центры для катализа биохимических реакций. Однако новейшие исследования, в том числе изучение механизмов процессинга РНК, привели к пересмотру этой точки зрения: оказалось, что дело обстоит не так категорично и вместе с тем не так просто.

В современной биологии термин **рибозим** стал общеупотребительным: им обозначают молекулу РНК со свойствами катализатора. Реакций *in vivo*, ведомых рибозимами, описано уже довольно много. В ряде случаев такие реакции не выходят за пределы одной молекулы РНК, которая, таким образом, сама создает себе условия для изменения своей структуры. Это так называемые внутримолекулярные, или автокаталитические, реакции; их катализ сводится к одному-единственному акту, в котором молекула-катализатор меняет свою собственную сущность.

В качестве примера можно привести интроны групп I и II, которые обладают способностью к вырезанию самих себя из молекулы (предшественника мРНК), в которой они содержатся. Используя интроны группы I в создании искусственных конструкций, удалось получить молекулы РНК, обладающие еще несколькими каталитическими активностями, производными от исходной активности.

Фермент рибонуклеаза Р представляет собой нуклеопротеин, в котором содержится всего одна молекула РНК. Именно эта молекула обладает каталитическими свойствами, необходимыми для расщепления субстрата, тРНК; остальная, белковая, часть фермента лишь усиливает эту способность (скорее всего, за счет поддержания оптимальной структуры каталитической РНК).

Обобщая характеристики двух реакций (приведенных в качестве примеров), можно утверждать, что существуют РНК, способные проводить внутримолекулярные или межмолекулярные реакции *in vitro* с образованием или расщеплением фосфодиэфирных связей. Все же, несмотря на то что специфичность реакции и основной катализ в таких случаях обеспечиваются РНК, для эффективного течения реакции *in vivo* может быть исключительно важна поддержка белков, ассоциированных с этой РНК.

Надо сказать, что сплайсинг предшественника мРНК — это лишь один из нескольких способов внесения изменений в структуру этой информационной молекулы. Так, существует способ, **редактирование РНК**, при котором происходит изменение азотистых оснований (или добавление рибонуклеотидов) по конкретным положениям предшественника. Добавление единичных рибонуклеотидов (чаще всего уридинов) характерно для процессинга транскриптов митохондриальных генов у некоторых низших эукариот. Подобно сплайсингу, такие добавления (вставки) происходят с разрывом и воссоединением межнуклеотидных связей; кроме того, для них нужен приток *информации* (в виде матрицы, или образца, новой, результирующей, последовательности).

27.2 Интроны группы I самоустраняются по механизму трансэтерификации

Ключевые понятия:

- Для автокаталитического сплайсинга *in vitro* интронам группы I требуются одновалентный катион, двухвалентный катион и нуклеотид гуанин.
- Весь сплайсинг ограничивается двумя реакциями трансэтерификации и обходится без дополнительных затрат энергии.
- В начале первой реакции трансэтерификации 3′-OH-конец кофактора атакует 5′-конец интрона.
- В начале второй реакции трансэтерификации свободный 3′-OH-конец, образовавшийся у первого экзона в результате первой реакции, атакует стык между этим экзоном и вторым интроном.
- В результате второй реакции интрон высвобождается как линейная молекула; она замыкается в кольцо после того, как ее 3′-OH-конец атакует одну из (двух возможных) внутренних межнуклеотидных связей.
- Внутренняя связь G^{414}–A^{16} может быть атакована другими молекулами; последующая реакция квалифицируется как *транс*-сплайсинг.

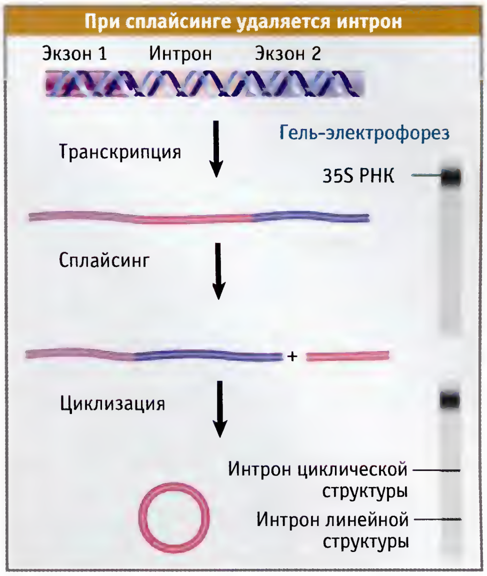


РИС. 27.1 Сплайсинг предшественника 35S рРНК *Tetrahymena* может быть оценен с помощью гель-электрофореза. Об удалении интрона свидетельствует появление на геле быстро движущейся легкой полосы. Интрон, имеющий циклическую структуру, при электрофорезе движется медленнее, что и демонстрирует наличие полосы, расположенной выше

Интроны группы I можно встретить в разных геномах. У низших эукариот (таких как инфузория *Tetrahymena thermophila* или плесневый гриб *Physarum polycephalum*) они есть в ядерных генах рибосомальной РНК. В митохондриальных генах грибов таких интронов попадается особенно много; их можно встретить у некоторых бактерий и даже в трех генах фага T4.

Все интроны этой группы умеют, выражаясь тривиально, вырезать самих себя из РНК. Этот процесс известен как автокаталитический сплайсинг, **автосплайсинг**. Способность с автокаталитическому сплайсингу есть и у интронов группы II, которым посвящен разд. 26.11 «Автосплайсинг интронов группы II происходит при помощи лассоа».

Автокаталитический сплайсинг был открыт у транскриптов генов рРНК *T. thermophila*. Гены обеих больших рРНК этой инфузории имеют типовую организацию; оба гена экспрессируются в составе одной большой транскрипционной единицы. Продукт транскрипции, 35S РНК, представляет собой молекулу-предшественник, 5′-концевая область которого содержит последовательность меньшей рРНК; от нее, в 3′-конец, идет последовательность большей (26S) рРНК. В некоторых подвидах *T. thermophila* последовательность 26S рРНК, в гене и первичном транскрипте, прервана одним коротким интроном.

Во время инкубации *in vitro* предшественник, 35S РНК, претерпевает самопроизвольный сплайсинг. При этом в реакционной среде начинает накапливаться короткий вырезанный интрон. Сначала он имеет вид линейного фрагмента длиной 400 оснований, затем этот фрагмент переходит в кольцевую форму. Схема событий приведена на РИС. 27.1.

Для самопроизвольного сплайсинга в среде кроме самой РНК должны присутствовать лишь одно- и двухвалентные катионы, а также кофактор нуклеотид гуанин (рибонуклеотиды с другим азотистым основанием, не G, не могут служить кофакторами этой реакции). Число фосфатов, которые есть у гуанозиннуклеотида, значения не имеет: ГТФ, ГДФ, ГМФ и даже просто гуанозин в принципе одинаково хорошо подходят для этой реакции (которая не требует затрат энергии). Очень важным требованием к гуанозиннуклеотиду является наличие у него свободной 3′-OH группы.

За судьбой гуанозиннуклеотида можно проследить при помощи радиоактивного мечения. После завершения сплайсинга метка обнаруживается в вырезанном линейном фрагменте интрона: бывший кофактор (остаток G) прикреплен к его 5′-концу посредством обычной фосфодиэфирной связи.

На РИС. 27.2 изображены три реакции переноса, из которых состоит автокаталитический сплайсинг. В первой реакции участвует гуанозиннуклеотид: он предоставляет свою свободную 3′-OH группу, которая атакует 5′-конец интрона. В этой реакции образуется связь между интроном и остатком G (бывшим кофактором); свободная 3′-OH-группа возникает при этом на конце первого экзона. Второй перенос похож на первый: свободная 3′-OH группа первого экзона атакует второй экзон. По всей видимости, первые два переноса сопряжены друг с другом, поскольку никаких свободных экзонов в природе не наблюдается. На самом деле, первые два переноса можно рассматривать как одну реакцию *лигирования* экзонов с высвобождением интрона. Теперь интрон имеет статус свободной молекулы, которая должна быть переведена в кольцевую форму. Для этого нужна третья реакция переноса.

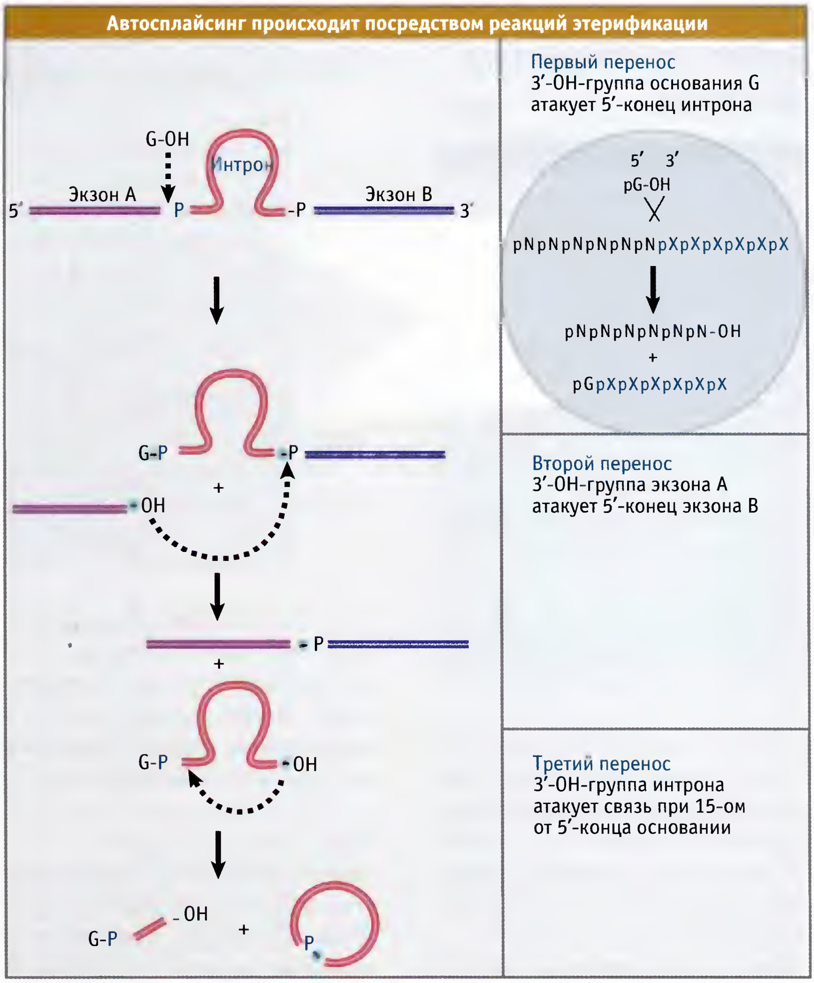


РИС. 27.2 Автосплайсинг реализуется с помощью реакций этерификации, в которых обмен связями происходит напрямую. Связи, образующиеся на каждой стадии, обозначены затемненными участками

Механизм у всех трех реакций один и тот же — трансэтерификация, при которой один фосфатный эфир переводится в другой фосфатный эфир без какого-либо промежуточного гидролиза. Обмен связями происходит напрямую; вся энергия при этом сохраняется, поэтому реакция не требует дополнительного гидролиза макроэргических связей (ни АТФ, ни ГТФ не расходуются). Если так, возникает вопрос: почему при постоянном энергетическом потенциале системы вместо того, чтобы установиться равновесию между продуктом и несплайсированным предшественником, реакция сплайсинга неизменно доходит до конца? Дело в том, что в физиологических условиях концентрация кофактора (ГТФ) гораздо выше, чем концентрация субстрата (РНК); это подталкивает весь процесс вперед, а изменения во вторичной структуре РНК придают реакциям необратимость.

Существование системы *in vitro*, где нет белков, а сплайсинг, тем не менее, успешно идет, дает полное *право признать за РНК способность к сплайсингу самой себя*. Во время сплайсинга молекула РНК принимает совершенно особую вторичную и третичную структуры, в которых функционально значимые химические группы сближены. Так формируется сайт связывания гуанозин нуклеотида; именно это связывание запускает реакции разрыва и воссоединения химических связей, показанные на рис. 27.2. По результатам экспериментов *in vitro* сплайсинг, безусловно, можно квалифицировать как свойство самой РНК; тем не менее в условиях *in vivo* все реакции проходят при участии белков, стабилизирующих структуру РНК.

Способность РНК к участию в реакциях переноса, из которых состоит сплайсинг, записана в самой нуклеотидной последовательности интрона. Поэтому вырезанный линейный интрон продолжает быть реакционноспособным; это показано на РИС. 27.3.

Вырезанный линейный интрон может перейти в кольцевую форму после атаки 3′-концевого остатка G (*противоположного* бывшему кофактору, находящемуся на 5′-конце) по одному из двух внутренних положений возле 5′-конца. При этом одна из внутренних межнуклеотидных связей рвется, и новый 5′-конец переносится на атаковавший 3′-ОН-конец интрона. Чаще

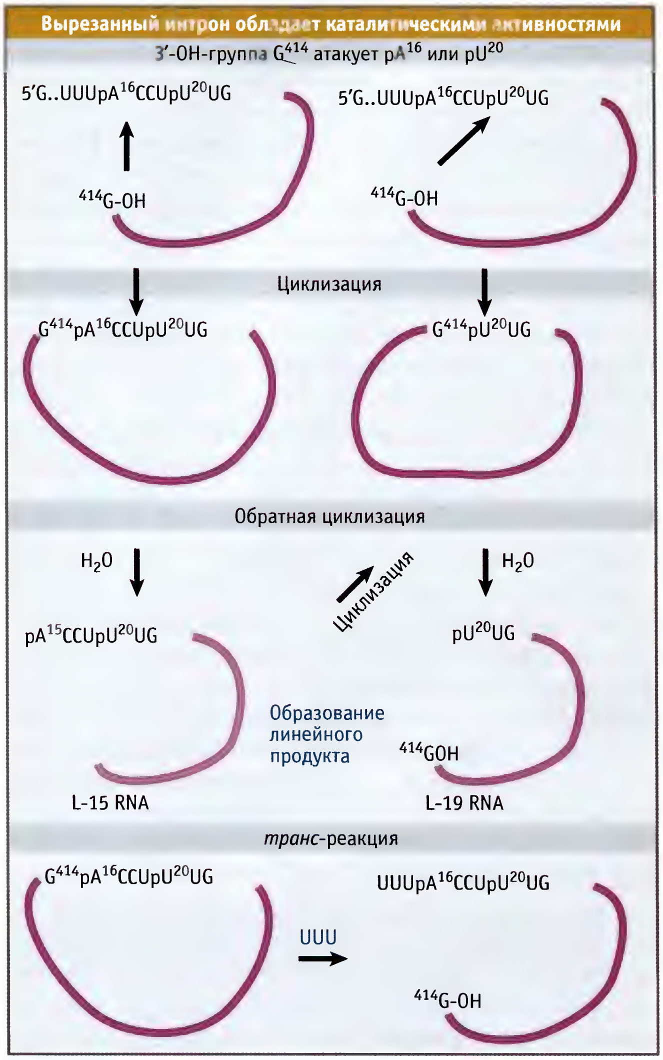


РИС. 27.3 Вырезанный итрон может формировать циклическую структуру, используя для взаимодействия с 5′-концом два внутренних сайта. Линейная структура интрона восстанавливается в результате реакции с водой или олигонуклеотидами

всего в этом *первичном замыкании* участвуют рибонуклеотиды G^{414} и A^{16}; реакция между ними изображена на рис. 27.2 как третий перенос. Иногда тот же G^{414} вместо A^{16} атакует U^{20}. В любом случае, первичное замыкание рождает кольцевой интрон, а также короткий линейный фрагмент, соответствующий 5′-концевому сегменту вырезанного линейного интрона. Длина короткого фрагмента составляет 15 оснований (или 19 оснований, если атака пришлась на U^{20}). На 5′-конце этого фрагмента находится гуанозиннуклеотид, служивший кофактором в первом переносе.

Далее, в условиях *in vitro*, кольцо (любой из двух вариантов) может самопроизвольно разомкнуться через прицельный гидролиз той самой связи, которой оно перед этим было замкнуто (G^{414}–A^{16} либо G^{414}–U^{20}). *Размыкание* кольца, замкнутого по A^{16}, дает реакционноспособную линейную молекулу, которая может совершить *вторичное замыкание*; для этого нужна атака G^{414} на U^{20}.

Конечным продуктом самопроизвольных реакций, происходящих с интроном после его выхода из состава предшественника, является РНК L-19 — это линейная молекула, полученная в результате размыкания более короткой кольцевой формы. Она обладает собственной энзиматической активностью, позволяющей ей катализировать наращивание коротких олигонуклеотидов. (Эта активность никак не помечена на рисунках к этому разделу, но см. рис. 27.8.)

Реакционная способность вырезанного интрона превосходит простое умение поворачивать вспять реакцию замыкания. Добавление в реакционную среду олигонуклеотида UUU провоцирует вскрытие первичного кольца. UUU, похожий на 3′-концевой край пятнадцатичленника, отделяемого при первичном замыкании, становится 5′-концевым краем линейной молекулы. Это уже *межмолекулярная* реакция, при которой две разные молекулы РНК объединяются в составе одной молекулы.

На примере серии реакций предшественника рРНК *T. thermophila* было показано, что автокаталитическая активность РНК проистекает из способности некоторых молекул РНК к образованию активного центра связывания кофактора, а также к реорганизации и сближению реакционноспособных групп в конформации, подходящей для нарушения и восстановления целостности молекул. Другие интроны группы I, хоть они и не были изучены так же хорошо, как интрон *T. thermophila*, по-видимому, обладают похожими свойствами. Таким образом, автокаталитический сплайсинг *in vitro* полностью обеспечивается самой РНК. Насколько же актуально участие белков в этом событии *in vivo*?

Интересные подтверждения участия белков в сплайсинге РНК были получены в работах с митохондриальными системами, где для сплайсинга интронов группы I нужны (*транс*-действующие) продукты других генов. Один удивительный пример участия белков в сплайсинге стал известен из наблюдений за мутантами *cyt18* гриба *Neurospora crassa*; в митохондриях этих мутантов не проходит сплайсинг интронов группы I. Выяснилось, что ген *cyt18* кодирует не что иное, как митохондриальную тирозил-тРНК-синтетазу. Вероятное объяснение фенотипического эффекта звучит так: успех сплайсинга зависит от принятия интроном такой же третичной структуры, как у тРНК; именно такая структура необходима для катализа трансэтерификации; должно быть, такая структура стабилизируется тирозил-тРНК-синтетазой.

Во всяком случае, очень интересен сам факт, что тРНК-синтетаза имеет явное отношение к сплайсингу. Этот пример говорит о том, что сплайсинг довольно долго эволюционировал как система реакций, опосредованных молекулами РНК, и лишь позднее их заменили РНК-связывающие белки, у которых прежде были какие-то другие функции. Такая популярная имитация, как автокаталитический сплайсинг *in vitro*, передает лишь биохимическую сущность события. Активный центр, действительно, закладывается в структуре самой РНК; однако для эффективного катализа *in vivo* активному центру необходима поддержка целого комплекта различных белков.

27.3 Интроны группы I имеют характерную вторичную структуру

Ключевые понятия:

- Во вторичной структуре интронов группы I есть девять двухцепочечных участков.
- Каталитическая активность интрона заключена в ядре, содержащем двухцепочечные участки P3, P4, P6 и P7.
- Участки P4 и P7 образованы парами комплементарных консервативных последовательностей.
- Последовательность, смежная с P7, комплементарна 3′-концевому сегменту интрона, содержащему «реактивный» G.

Каждый интрон, относящийся к группе I, может быть организован во вторичную структуру с девятью двухцепочечными участками, обозначаемыми как P1, P2,... P9. Модель вторичной структуры интрона *T. thermophila* представлена на РИС. 27.4. Из девяти двухцепочечных участков, обнаруживаемых во вторичной структуре, два участка, P4 и P7, образованы комплементарными консервативными элементами, которые у всех интронов этой группы примерно одинаковы. Двойная спираль P4 образована комплементарными элементами P и Q; аналогичным образом, P7 образована элементами R и S. Остальные двухцепочечные участки вторичной структуры не имеют выраженного сходства с аналогичными участками других интронов. Анализ мутаций позволяет говорить о наличии у интрона функционально значимого ядра, в которое входят P3, P4 и P6; это минимальная область интрона, способная к выполнению катализа. Размеры интронов группы I очень вариабельны; консервативные участки находятся на значительном расстоянии от концов интрона (то есть от стыков между экзонами).

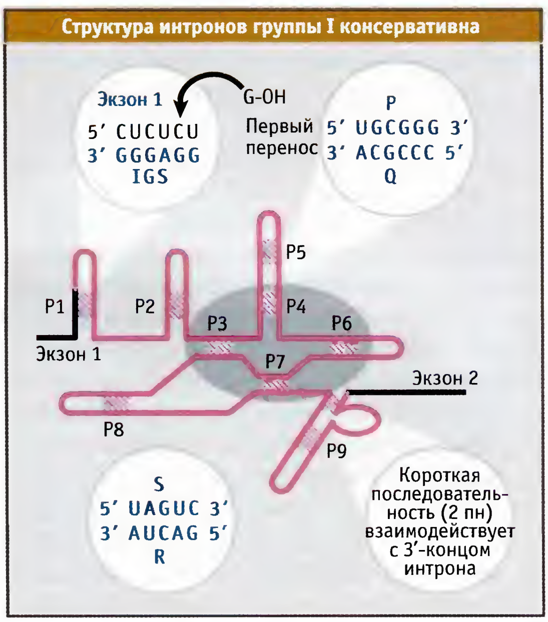


РИС. 27.4 **Интроны группы I имеют схожую вторичную структуру, формируемую девятью регионами спаривания. Регионы P4 и P7 консервативны и образованы последовательностями P, Q, R и S. P1 формируется в результате спаривания между концом левого экзона и ВНП (внутренней направляющей последовательностью) интрона; участок, расположенный между P7 и P9, спаривается с 3′-концом интрона**

Наличие двухцепочечных участков во вторичной структуре интрона помогает привести стыки между экзонами в конформацию, благоприятствующую сплайсингу. Участок P1 содержит край первого экзона (который на схемах изображен слева от интрона). Последовательность внутри интрона, взаимодействующую с краем первого экзона, называют внутренней наводящей последовательностью (ВНП). Одно время считалось, что участок интрона, следующий сразу за ВНП, обязательно должен взаимодействовать с комплементарным краем второго экзона, так что ВНП сводит вместе два экзона, и стык между ними *возникает* именно благодаря ВНП. (Впоследствии оказалось, что взаимодействие ВНП со вторым экзоном имеет место не во всех интронах этой группы и что для успеха сплайсинга оно не так уж и важно.) Наконец, один очень короткий внутренний сегмент интрона, длиной от 2 оснований, расположенный между участками P7 и P9, взаимодействует с сегментом, расположенным непосредственно перед G на 3′-конце интрона (у *T. thermophila* — перед G^{414}).

Чтобы интрон стал рибозимом, нужно, чтобы его участки взаимодействовали друг с другом в точном соответствии с первичной структурой интрона. Изменение рибонуклеотидной последовательности интрона может помешать сплайсингу. В отличие от фенотипа мутаций, имеющих место в митохондриях (где нарушения сплайсинга легко проявляются в условиях *in vivo*), фенотип мутаций в интроне *T. thermophila* лучше всего проявляется в бактериальной системе. На РИС. 27.5 изображена искусственная генно-инженерная конструкция; за ходом сплайсинга транскриптов этой конструкции можно следить, используя клетки *E. coli* в качестве реакционной среды. При создании этой конструкции интрон, подлежащий автокаталитическому сплайсингу, был помещен в положение, благодаря которому он прерывает кодирующую последовательность β-галактозидазы (по десятому кодону). Таким образом, синтез β-галактозидазы и появление в клетке *E. coli* цветного продукта активности этого фермента в данном случае становится возможным только после сплайсинга, т.е. после удаления интрона из мРНК.

Как оказалось, синтез β-галактозидазы в этой системе может идти весьма интенсивно; это говорит о том, что сплайсинг достаточно независим от внешних условий (хоть мы ничего и не знаем о возможной роли бактериальных белков в успехе этого сплайсинга, полноценный сплайсинг в бесклеточной системе *in vitro* говорит о том же самом).

Используя ту же систему, можно проверять, как влияют на ход сплайсинга различные мутации. Те мутации,

что меняют консенсус, общий для всех интронов группы I, и, как следствие, нарушают совершенство двухцепочечных участков, обычно приводят к полной остановке сплайсинга. Их эффект можно компенсировать за счет введения других мутаций, восстанавливающих комплементарность оснований в важнейших «узлах» вторичной структуры интрона.

В интронах митохондриальных генов аналогичные мутации характеризуются похожими фенотипическими эффектами. Мутация в одном консервативном элементе нуклеотидной последовательности интрона может быть компенсирована мутацией в комплементарном элементе, возвращающей интрону его правильную вторичную структуру. Так, мутации в консенсусе R компенсируются мутациями в консенсусе S.

Обобщая сказанное, можно утверждать, что автосплайсинг, проводимый интроном группы I, зависит от его особой вторичной структуры, основанной на взаимодействиях комплементарных консервативных элементов. Основной принцип, который здесь можно сформулировать, звучит так: рибонуклеотидные последовательности внутри интрона, далекие от стыков между экзонами, участвуют в образовании каталитических активных центров, организующих все реакции, из которых состоит сплайсинг.

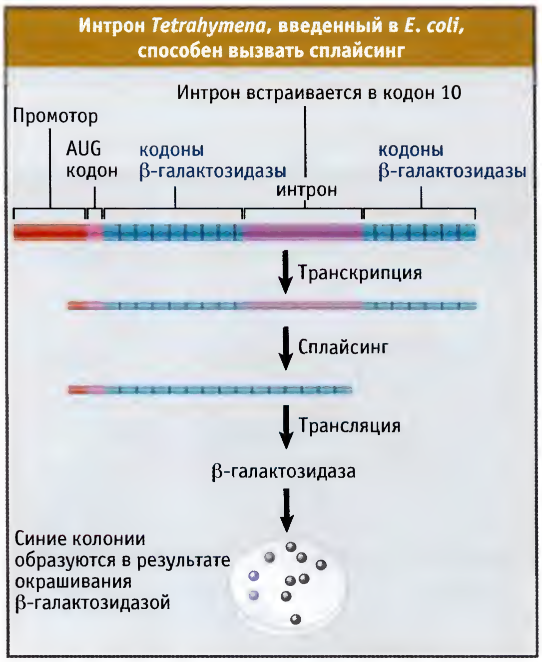


РИС. 27.5 Введение интрона *Tetrahymena* в кодирующую последовательность гена β-галактозидазы приводит к возникновению у *E. coli* способности к аутосплайсингу. Синтез β-галактозидазы может быть оценен с помощью добавления в культуральную среду компонентов, окрашивающихся ферментом в синий цвет. Последовательность вносится в *E. coli* бактериофагом, поэтому присутствие синих колоний (содержащих инфицированные бактерии) свидетельствует об удачном сплайсинге

27.4 Рибозимы проявляют различные каталитические активности

Ключевые понятия:

- У интрона группы I можно поменять специфичность к субстрату; для этого некоторые элементы его последовательности надо заменить альтернативными.
- Реакции сплайсинга имеют стандартную кинетику ферментативных реакций.
- Начальным этапом эволюции каталитических активностей РНК могли быть реакции, затрагивавшие связи 2′-OH.

Каталитическая активность интронов первой группы была открыта как их способность к автосплайсингу. Тем не менее в условиях *in vitro* такие интроны могут проявлять еще целый ряд каталитических активностей. Любая из этих активностей обслуживает какой-нибудь вариант реакции трансэтерификации. В этом разделе мы попробуем взглянуть на эти варианты с позиций их сходства с реакциями сплайсинга.

Любая каталитическая активность интрона первой группы определяется его способностью к образованию каталитически активных центров — функциональных эквивалентов активных центров ферментов белковой природы. На РИС. 27.6 снова изображен сплайсинг (та же серия реакций, что и на рис. 27.2); на этот раз сплайсинг показан с позиций загрузки активных центров рибозима.

Один из активных центров, так называемый сайт связывания субстрата, образован двойной спиралью P1, в которой 3′-конец первого экзона приходится на участок ВНП. Второй активный центр, сайт связывания гуанозина, образован другой двойной спиралью, P7. Сайт связывания гуанозина может быть занят либо кофактором, либо «реактивным» 3′-концевым остатком интрона, G (у *T. thermophila* это G^{414}). В первой реакции переноса там находится кофактор гуанозиннуклеотид, в последующих реакциях там находится G^{414}. Первые две реакции рождают соединение между экзонами. Третья реакция (у которой нет очевидного биологического смысла) рождает кольцевой интрон.

Связывание субстратов сопровождается значительными изменениями в структуре интрона. До связывания 5′-конец участка ВНП находится близко к P2 и P8; после связывания, когда ВНП уже участвует в образовании двойной спирали P1, его 5′-конец оказывается близко к консервативным рибонуклеотидным положениям между P4 и P5. Ход сплайсинга можно отслеживать по контактам, детектируемым во вторичной структуре интрона; эти контакты отмечены на РИС. 27.7. В третичной структуре интрона две точки, с которыми в альтернативном режиме контактирует двойная спираль P1, отстоят друг от друга на расстояние 37 Å; это означает, что в процессе работы рибозима двойная спираль P1 не стоит на месте, но совершает значительные перемещения.

РНК L-19 представляет собой продукт размыкания кольцевого интрона (на рис. 27.3 эта реакция изоб-

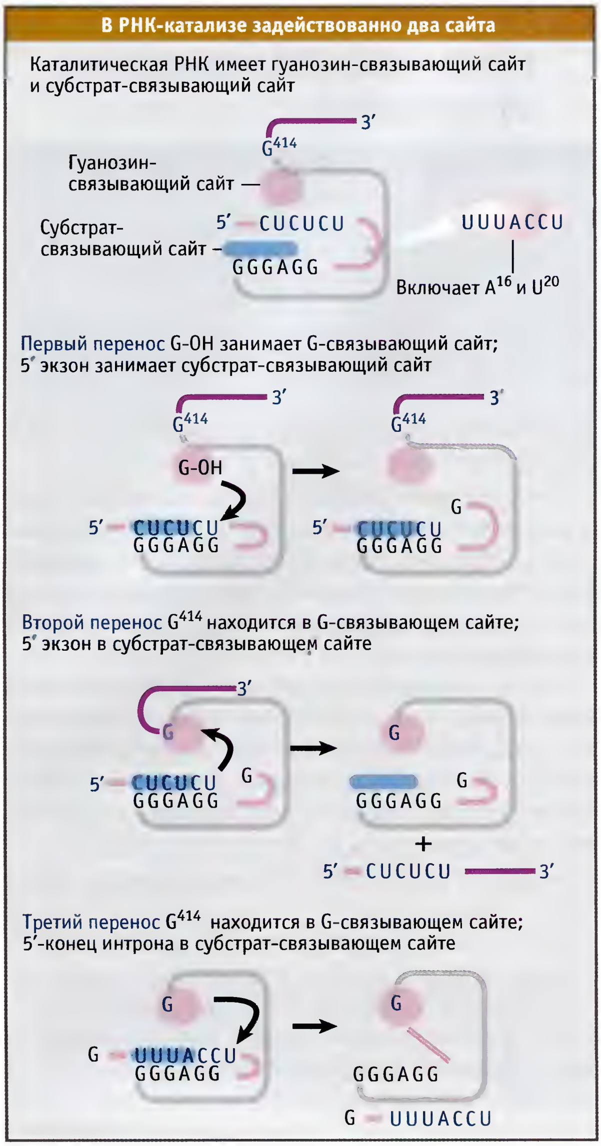


РИС. 27.6 Удаление интронов группы I из рРНК *Tetrahymena* происходит в результате реакций между оккупантами гуанозин-связывающего сайта и субстрат-связывающего сайта. Левый экзон обозначен розовым цветом, правый — пурпурным

ражена как последняя из внутримолекулярных перестроек, составляющих сплайсинг). Несмотря на свой вид, РНК L-19 все еще является рибозимом (хотя ее активность можно считать остаточной по отношению к полновесной каталитической активности, проявленной интактным интроном в реакциях сплайсинга). Эта остаточная активность определяется наличием у L-19 последовательности ВНП, комплементарной ВНП. Взаимодействие этих последовательностей происходит в активном центре связывания субстрата; второй активный центр рибозима отвечает за связывание концевого G^{414} или свободного гуанозина в G-связывающем сайте (то есть отвечает за атаку, с которой начинается любая реакция трансэтерификации).

На **РИС.** 27.8 показан механизм наращивания олигонуклеотида C_5 до C_6 в исполнении РНК L-19. Олигонуклеотид C_5 находится в сайте связывания субстрата; в это время в сайте связывания гуанозина находится G^{414}. Из этого исходного положения 3′-концевой цитидин C_5 переносится на G^{414}; затем он переносится обратно, в сайт связывания субстрата, но уже на новую молекулу C_5. Многократное повторение этой реакции приводит к накоплению в реакционной среде все более длинных олигоцитидинов C_n. Участие L-19 в этой реакции, безусловно, можно назвать катализом, поскольку сама L-19 при этом совершенно не меняется и может бесперебойно поддерживать серию из множества циклов. Таким образом, РНК L-19 однозначно квалифицируется как рибозим, обладающий активностью нуклеотидил-трансферазы.

Еще несколько реакций, поддерживаемых рибозимами, представлено на **РИС.** 27.9. Рибозим может вести себя как эндорибонуклеаза, специфичная к рибонуклеотидной последовательности, комплементарной ВНП. В приведенном примере ВНП связывает сегмент РНК, идентичный краю первого экзона (на самом деле этот сегмент, CUCU, принадлежит совсем другому, обособленному, субстрату — другой молекуле РНК). Для этой реакции нужен кофактор гуанозиннуклеотид; он занимает свой сайт и оттуда атакует последовательность CUCU (точно так же, как он атакует конец первого экзона в самом начале сплайсинга). В результате последовательность-мишень расщепляется на 5′-концевую область, с краем CUCU как у первого экзона, и 3′-концевую область, с добавленным G на 5′-конце. Поменять специфичность рибозима можно путем внесения

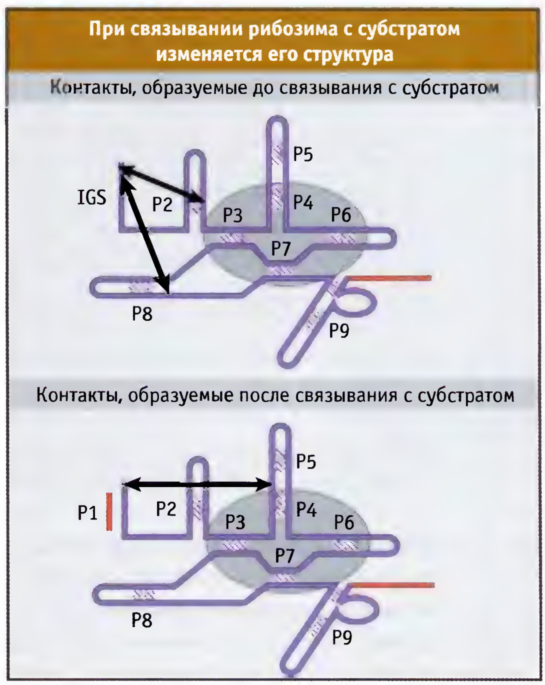


РИС. 27.7 Положение ВНП в третичной структуре изменяется, когда при связывании с субстратом формируется P1

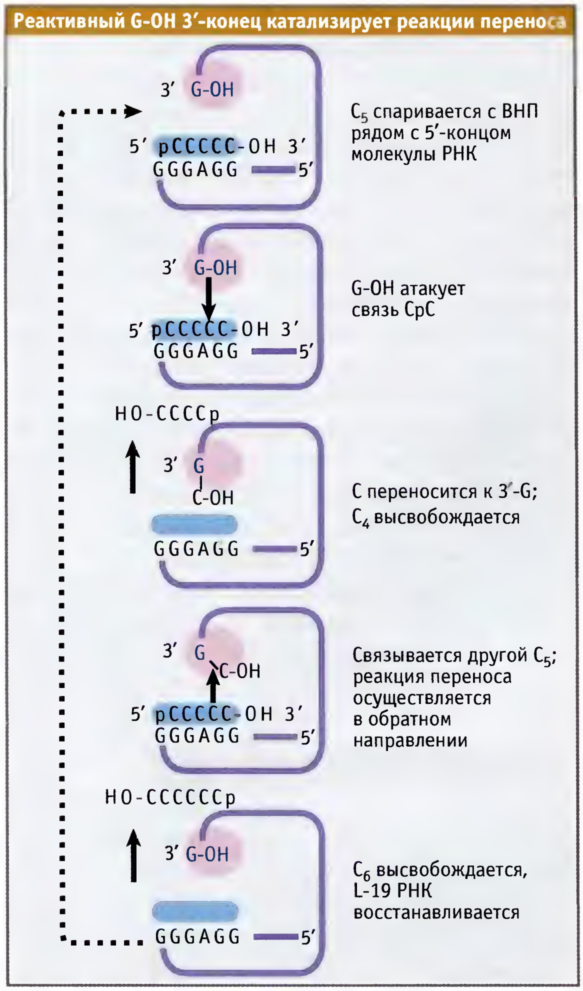


РИС. 27.8 Линейная РНК L-19 способна связывать C в субстрат-связывающем сайте; реактивный G-3′-OH-конец располагается в G-связывающем сайте и катализирует реакции переноса, в результате которых из двух C_5 олигонуклеотидов образуются C_4 и C_6 олигонуклеотиды

мутаций в элемент ВНП; рибозим будет узнавать последовательность, комплементарную новой версии ВНП.

Другой компетенцией рибозима является лигирование (т. е. соединение концов) двух молекул РНК. Для выполнения этой реакции одна молекула РНК со свободным 3′-OH-концом должна закрепиться в сайте связывания субстрата (специфичность связывания можно модулировать коррекцией последовательности ВНП). Во втором активном центре должна ждать вторая молекула РНК, обязательно с G на 5′-конце (потому что за этот G рибозим ее удерживает). В итоге атака свободного гидроксила на фосфатную связь, направленная из первого центра во второй, соединяет две молекулы в одну, с потерей G, бывшего 5′-концевого остатка второй молекулы.

Дефосфорилирование (так называемая фосфатазная реакция) в исполнении рибозима не является прямой имитацией какой-либо из реакций сплайсинга, но тоже может иметь место. Для этой реакции олигонуклеотид, комплементарный ВНП и оканчивающийся на 3′-фосфат, должен быть атакован G^{414}. В результате атаки фосфат переносится на G^{414} и оставляет олигонуклеотид со свободным 3′-OH-концом. После того как этот олигонуклеотид покинет сайт связывания субстрата, фосфат может быть перенесен обратно, из сайта связывания гуанозина в сайт связывания субстрата, либо на другой олигонуклеотид (что эквивалентно обратному ходу реакции), либо на воду (что эквивалентно выходу неорганического фосфата, которым всегда завершается *истинное* дефосфорилирование).

Реакции, где катализатором служит РНК, можно, подобно классическим ферментативным процессам, описать в терминах кинетики Михаэлиса-Ментен. Основные характеристики реакций сведены в таблицу на **РИС. 27.10**. Для всех реакций, в которых участвуют рибозимы, характерны низкие значения K_M; это говорит о специфичности каталитических РНК к своим субстратам. Низкие значения валовых показателей отражают низкую эффективность катализа. В целом каталитические молекулы РНК ведут себя как классические ферменты, хотя по сравнению с белками они все делают очень медленно (белки, в большинстве своем, совершают от 10^3 до 10^6 актов катализа в минуту).

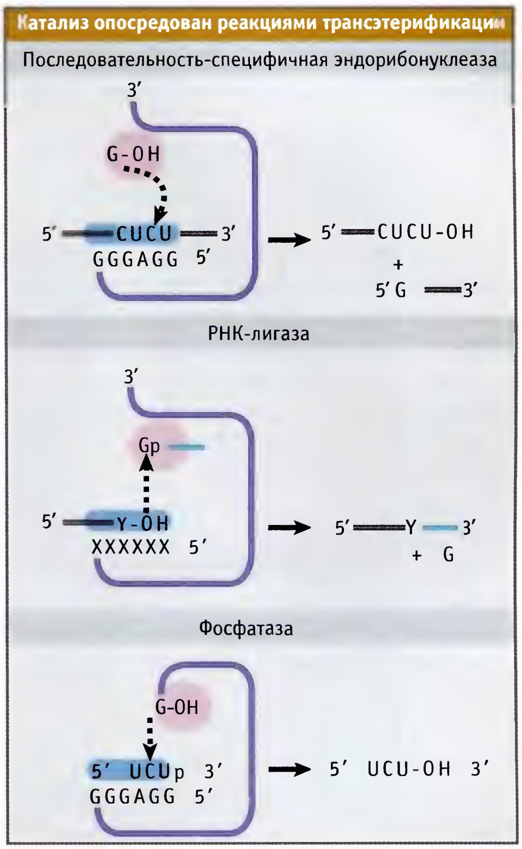


РИС. 27.9 Каталитическая реакция с участием рибозима осуществляется с помощью реакций трансэтерификации между группами, расположенными в субстрат-связывающем сайте и группами, расположенными в G-связывающем сайте

РНК-катализ является энзиматическим

Фермент	Субстрат	K_M (мМ)	Число оборотов (в мин)
19 оснований вирусной РНК	24-нуклеотидная РНК	0,0006	0,5
Интрон L-19	CCCCCC	0,04	1,7
РНК РНКазы Р	пре-тРНК	0,00003	0,4
РНКаза Р	пре-тРНК	0,00003	29
РНКаза Т1	GpA	0,05	5 700
β-галактозидаза	лактоза	4,0	12 500

РИС. 27.10 Реакция, катализируемая РНК, описывается теми же параметрами, что и реакция, катализируемая белками, отличаясь только в том, что протекает медленнее. K_M характеризует такую концентрацию субстрата, при которой скорость реакции достигает половины своего максимума; это величина обратная аффинности фермента к субстрату. Число оборотов отражает число молекул субстрата, переработанных одним каталитическим сайтом в единицу времени

Обнаружение того факта, что каталитические РНК могут регулироваться от взаимодействий с небольшими лигандами, перевернуло новую страницу в представлениях об активности рибозимов (см. разд. 13.7 «Регуляция трансляции с помощью малых молекул РНК»). Один из вариантов такой регуляции (**riboswitch**) изображен на РИС. 27.11. Суть его в следующем: небольшая молекула, метаболит $GlcN_6P$, связывается с рибозимом и побуждает его к внутримолекулярной реакции расщепления РНК (то есть лиганд заставляет рибозим совершить автокаталитическое расщепление, проще говоря, разрезать самого себя). Так устроена саморегуляция продукции $GlcN_6P$ по типу отрицательной обратной связи — ведь рибозим есть не что иное, как 5'-концевая нетранслируемая область мРНК, кодирующая ключевой фермент продукции $GlcN_6P$. Автокаталитическое расщепление этой мРНК предотвращает ее трансляцию.

Теперь зададим себе вопрос: каким образом молекуле РНК удается обеспечивать каталитически актив-

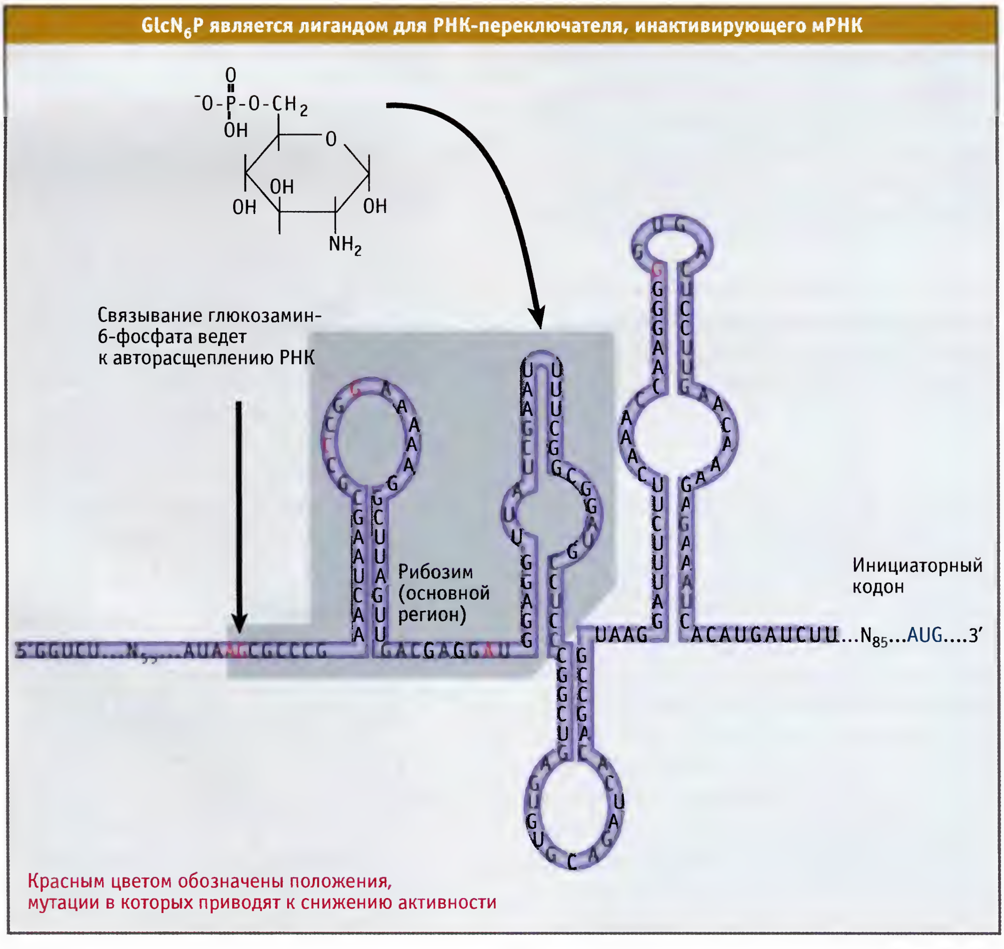


РИС. 27.11 Рибозим расположен внутри 5'-нетранслируемого региона мРНК, которая кодирует фермент, участвующий в образовании глюкозамин-6-фосфата. Когда глюкозамин-6-фосфат связывается с рибозимом, тот отщепляет 5'-конец мРНК, предотвращая тем самым дальнейшую продукцию фермента

ный центр? Активный центр можно представить себе как поверхность, образованную активными химическими группами, пребывающими в фиксированных взаимоотношениях друг с другом и с субстратом. В активных центрах различных белков роль активных групп играют радикалы аминокислотных остатков; в своей совокупности эти радикалы очень разнообразны, имеют много положительно и отрицательно заряженных ионных групп, а также гидрофобных групп, и представляют собой богатейший материал для создания активных центров. У РНК в качестве активных центров служат, по большей части, доступные сочетания азотистых оснований. Поддержание этих сочетаний в определенных положениях может осуществляться за счет жесткости вторичной (или третичной) структуры молекулы РНК. Оказывается, даже этих ограниченных возможностей достаточно для создания фронта химически активных групп в конфигурации, которая обеспечивает среду для изменения паттерна химических связей у субстрата. Взаимодействие РНК-катализатора с РНК-субстратом неизбежно сопровождается взаимодействием их комплементарных участков, обеспечивающим реакционное окружение. Очень важную роль в РНК-катализе играют двухвалентные катионы (обычно это Mg^{2+}), они присутствуют в активном центре и координируют и фиксируют позиции различных активных химических групп. Двухвалентные катионы принимают особенное участие в эндонуклеолитической активности рибозимов некоторых вирусоидов (см. разд. 27.9 «Вироиды обладают каталитической активностью»).

Открытие у РНК каталитической активности дало мощный толчок развитию теории эволюции. Явный раскол генетического аппарата (во всех компонентах которого присутствует нуклеиновая кислота, но при этом весь катализ отводится белкам) озадачивал исследователей во все времена. Всегда казалось маловероятным, чтобы в самых первых реплицирующихся системах (т. е. живых системах) могли одновременно присутствовать и белки, и нуклеиновые кислоты. Но предположим, что первые живые системы содержали одну лишь самореплицирующуюся нуклеиновую кислоту, обладавшую каталитическими активностями — очень примитивными, но вполне достаточными для того, чтобы разрывать и устанавливать фосфодиэфирные связи. Если предположить, что участие связей 2′-ОН в реакциях сплайсинга у современных организмов является отголоском примитивного древнего катализа, придется признать, что, вероятнее всего, первой генетической матрицей на Земле служила не ДНК, а РНК. (У ДНК отсутствует 2′-ОН-группа.) Продолжая это рассуждение, можно представить, что белки смогли завоевать себе место в живых системах благодаря своей способности к стабилизации структуры РНК. Постепенно белки, благодаря своей непревзойденной структурной пластичности, фактически монополизировали катализ, и уже на этой основе строился сложнейший аппарат экспрессии генов современных живых организмов.

27.5 Некоторые интроны группы I кодируют эндонуклеазы, поддерживающие их мобильность

Ключевые понятия:

- Мобильные интроны способны внедряться в новые сайты.
- Мобильные интроны группы I кодируют эндонуклеазы, необходимые для их внедрения.
- Внедрение интрона в сайт-мишень происходит по всем правилам репликативной транспозиции.

Некоторые интроны групп I и II содержат открытые рамки считывания, которые транслируются как функциональные белки. Экспрессия генетического материала, содержащегося в таких интронах, придает им *мобильность*. Это значит, что при поддержке продукта экспрессии ДНК-версия интрона (оригинальная геномная копия или обратная копия с транскрипта) может встроиться в новый сайт геномной ДНК. Интроны групп I и II распространены чрезвычайно широко; они представлены во всех царствах живых организмов. Перемещение интронов группы I опосредовано ДНК; перемещение интронов группы II опосредовано РНК.

Сам факт мобильности интронов был впервые установлен в генетических экспериментах. Обнаружилось, что в некоторых локусах аллели различаются наличием или отсутствием интрона. Особенно часто такой полиморфизм можно встретить в митохондриях грибов. Наличие полиморфизма в пределах одной клетки соответствует предположению о том, что эти интроны *появляются* как вставки генетического материала внутрь гена. Механизм их появления может быть приоткрыт анализом рекомбинации, затрагивающей ген больших рРНК в митохондриях гибридных дрожжевых клеток.

В гене больших рРНК есть интрон группы I, несущий кодирующую последовательность. У некоторых штаммов дрожжей (ω^+) этот интрон присутствует, а у некоторых (ω^-) он отсутствует. Скрещивания двух разных штаммов, на языке генетики, *полярны*: все потомство обладает генотипом ω^+. Для объяснения полярности проще всего предположить, что ω^+ играет в скрещивании роль донора генетического материала, а ω^- — роль реципиента; в результате в геноме ω^- у гибридов $\omega^+ \times \omega^-$ возникает новая копия интрона, и все потомство обретает генотип ω^+.

Полярность наследования может быть нарушена мутациями в митохондриальном геноме одного из родительских штаммов. Такие мутации нормализуют сегрегацию генотипов; в результате потомство содержит равные пропорции клеток ω^+ и ω^-. Анализ нуклеотидных положений, на которые приходятся такие мутации, объясняет природу полярности. Мутации в штамме ω^- возникают близко к сайту внедрения интрона. Мутации в штамме ω^+ приходятся на открытую рамку считывания, которая содержится в интроне, и препятствуют ее экспрессии. На основании данного анализа была построена

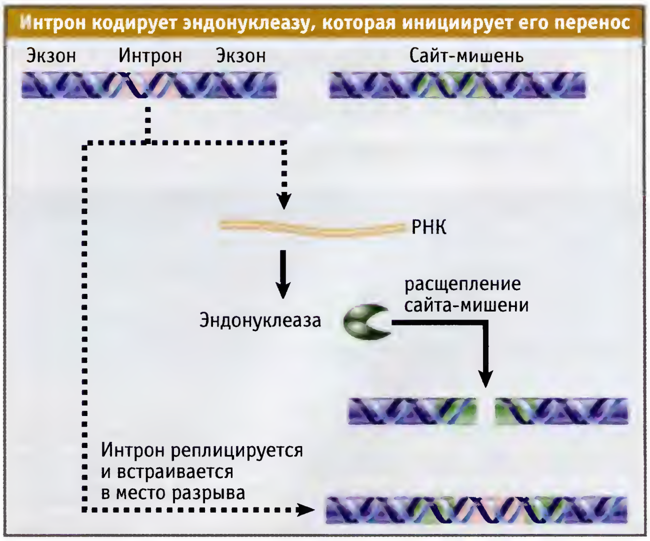


РИС. 27.12 Интрон кодирует эндонуклеазу, которая вносит разрыв в двухцепочечную молекулу ДНК. Последовательность интрона дуплицируется и встраивается в место разрыва

модель, изображенная на РИС. 27.12: белок, кодируемый интроном в штамме ω^+, узнает сайт внедрения интрона в мтДНК штамма ω^-; в результате у потомства происходит экспансия интрона, наблюдаемая как полярность скрещивания в пользу ω^+.

Каким действием должен обладать белок, продукт интрона ω? Он функционирует как эндонуклеаза: *узнает конкретную последовательность-мишень внутри гена* ω^- *и вносит в нее двухцепочечный разрыв* (в форме двух надрезов). Последовательность-мишень имеет размер 18 пн и содержит сайт внедрения интрона. Эндонуклеаза надрезает последовательность-мишень по обеим цепям ДНК; каждый надрез отстоит от сайта внедрения на два нуклеотида, в направлении 3'. Два надреза, разнесенные на расстояние 4 пн, эквивалентны двухцепочечному разрыву в ДНК. В этот разрыв с обеих сторон выступают четырехнуклеотидные одноцепочечные участки.

Такие прицельные надрезы, приводящие к разрывам с «липкими» краями, обычно сопровождают перемещения транспозонов (мобильных элементов генома; см. гл. 21 «Транспозоны»). В данном случае разрыв в ДНК инициирует процесс конверсии гена (процесс замены последовательности ω^- дубликатом последовательности ω^+). Этот процесс можно классифицировать как разновидность репликативной транспозиции, которая происходит без выхода за пределы уровня ДНК. Внедрение интрона в мишень, узнаваемую эндонуклеазой, аннулирует ее специфичность и придает всему гену устойчивость к этой эндонуклеазе, т. е. стабильность.

Многие интроны группы I кодируют эндонуклеазы, обеспечивающие им мобильность. Известно уже несколько семейств таких эндонуклеаз, но у всех у них есть одна общая черта: рядом с активным центром располагается аминокислотный мотив LAGLIDADG. Нередко бывает, что похожие интроны кодируют различающиеся эндонуклеазы. Механизмы внедрения разных интронов отличаются деталями; например, эндонуклеаза, кодируемая интроном *td* фага T4, режет последовательность-мишень, расположенную на 24 пн выше сайта внедрения интрона (т.е. конверсия стартует оттуда). Здесь можно добавить, что похожие последовательности-мишени есть у интеинов (сегментов полипептида, подлежащих автокаталитическому сплайсингу). Дело в том, что некоторые интеины «по совместительству» работают эндонуклеазами (см. разд. 27.12 «Сплайсинг белка — автокаталитический процесс»).

Многообразие эндонуклеаз, кодируемых интронами группы I, означает, что их мишени не похожи между собой. Это одни из самых протяженных и, следовательно, одни из самых специфичных мишеней, которые известны для эндонуклеаз; их размеры составляют от 14 до 40 пн. Высокая специфичность узнавания гарантирует, что интрон будет встроен в предназначенную для него мишень и ни в какое другое место генома. Этот эффект известен как **хоминг интронов**; обеспечением данного эффекта служит специфичность эндонуклеаз, известных как **хоминг-эндонуклеазы**.

Интроны, содержащие кодирующие последовательности для эндонуклеаз, были обнаружены у множества бактерий и низших эукариот. Это подкрепляет представление о том, что интроны, которые содержат кодирующие последовательности, возникали как независимые элементы генома.

27.6 Интроны группы II могут кодировать полифункциональные белки

Ключевые понятия:

- Интроны группы II способны к автокаталитическому сплайсингу *in vitro*; в условиях *in vivo* им обычно содействуют белки, которые кодирует сам интрон.
- Единственная находящаяся в интроне открытая рамка считывания дает белок с активностями обратной транскриптазы, матуразы и эндонуклеазы; также в этом белке есть ДНК-связывающий мотив.
- Обратная транскриптаза синтезирует ДНК-копию интрона на матрице РНК; эта копия ведет себя подобно копиям ретропозонов.
- Внося надрезы по обеим цепям ДНК, эндонуклеаза готовит сайт-мишень к внедрению интрона.

Самые изученные мобильные интроны группы II кодируют по одному-единственному белку; открытая рамка считывания находится за пределами каталитического ядра интрона. В белке, типичном продукте трансляции этой рамки, есть несколько доменов с самостоятельными функциями. Один из доменов, образованный N-концевой частью полипептида, имеет активность обратной транскриптазы. Активность центрального домена направлена на придание интрону каталитически активной

структуры (это так называемая активность матуразы; о ней см. разд. 27.7 «Некоторым каталитическим интронам требуется матураза»). Кроме того, в этом белке есть ДНК-связывающий мотив, а также С-концевой домен с активностью эндонуклеазы.

Эндонуклеаза инициирует реакцию транспозиции; она играет в хоминге ту же роль, что и аналогичные эндонуклеазы интронов группы I (см. разд. 27.5 «Некоторые интроны группы I кодируют эндонуклеазы, поддерживающие их мобильность»). Обратная транскриптаза синтезирует ДНК-копию интрона, подлежащую внедрению в свой сайт (по всем правилам хоминга). Иногда эндонуклеаза все-таки режет мишени, похожие на истинную мишень, но не идентичные ей, и находящиеся в другом месте. Это приводит к ошибочному внедрению интрона в несоответствующие последовательности; в таких случаях поведение интрона напоминает поведение большинства транспозонов, для которых характерна более низкая специфичность внедрения.

На РИС. 27.13 изображено событие транспозиции интрона группы II. Сначала эндонуклеаза вносит двухцепочечный разрыв в сайт-мишень. Надо сказать, что в данном случае активность эндонуклеазы является плодом совместных усилий соответствующего домена белка и РНК интрона. Домен белка перерезает антисмысловую цепь ДНК, а РНК интрона фактически перерезает смысловую цепь. Эти два надреза подготавливают сайт-мишень к внедрению интрона.

РНК-последовательность интрона служит матрицей для синтеза кДНК. Почти все белковые продукты интронов группы II обладают активностью обратной транскриптазы, специфичной к последовательности интрона. После того как обратная транскриптаза синтезирует копию интрона, внедрение можно считать состоявшимся: теперь интрон, в форме двухцепочечной ДНК, занимает надлежащий сайт генома.

Механизм внедрения мобильных интронов группы II напоминает транспозицию ретровирусов, где промежуточным продуктом обязательно выступает РНК (см. разд. 22.2 «Жизненный цикл ретровируса включает в себя стадии, подобные транспозиции»). Ретротранспозиция интронов группы II больше всего напоминает ретротранспозицию ретропозонов, у которых отсутствуют ДКП. Они образуют свободную 3′-ОН-группу, с которой, после внесения надреза в сайт-мишень, стартует синтез ДНК (см. рис. 22.20 к разд. 22.12 «LINES использует собственную эндонуклеазу, чтобы создать затравочный конец для синтеза ДНК»).

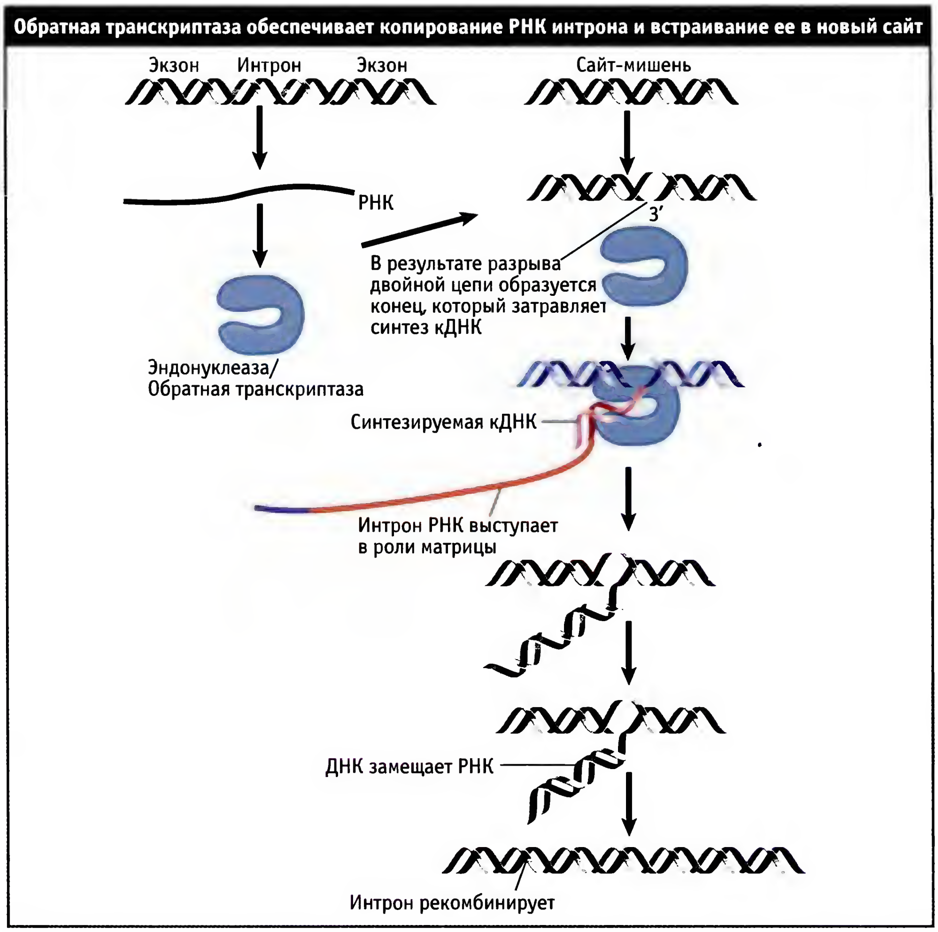


РИС. 27.13 Обратная транскриптаза, кодируемая интроном, обеспечивает синтез кДНК с матрицы РНК и встраивание ее в сайт-мишень, который образуется при разрыве двойной цепи ДНК

27.7 Некоторым автокаталитическим интронам требуется матураза

Ключевые понятия:

- Матураза, кодируемая самим интроном, помогает ему принять каталитически активную вторичную и третичную структуру.

Как известно, интроны групп I и II способны к автокаталитическому сплайсингу *in vitro*. Но в физиологических условиях им требуется поддержка белков. Интроны обеих групп могут кодировать белок, который имеет активность **матуразы** и помогает своему интрону в проведении реакций сплайсинга.

Активность матуразы принадлежит определенной части полипептида, продукта трансляции открытой рамки считывания, которую содержит интрон. Активность матуразы обнаружена у многих хоминг-эндонуклеаз (см. разд. 27.5 «Некоторые интроны группы I кодируют эндонуклеазы, поддерживающие их мобильность»). О независимости активностей эндонуклеазы и матуразы, сочетаемых в одном белке, свидетельствуют данные анализа мутаций.

Согласно данным структурных исследований, активности эндонуклеазы и матуразы обеспечиваются разными участками одного белка, причем каждый из них находится в отдельном домене. Активный центр эндонуклеазы обладает сродством к ДНК, в то время как активный центр матуразы связывает интрон в РНК. На РИС. 27.14 изображена структура такого белка в комплексе с ДНК. На рисунке видно, что в домене со свойствами эндонуклеазы (именно этот домен связывает ДНК) можно вычленить параллельные α-спирали, содержащие отличительный мотив LAGLIDADG, предшествующий активному центру эндонуклеазы. Активный центр матуразы



РИС. 27.14 Способный к хомингу интрон кодирует эндонуклеазу семейства LAGLIDADG, обладающую матуразной активностью. Последовательность LAGLIDADG входит в состав двух α-спиралей, заканчивающихся каталитическими аминокислотами в непосредственной близости от ДНК-дуплекса. Активный центр матуразы представляет собой аргининовый остаток, расположенный на поверхности белка

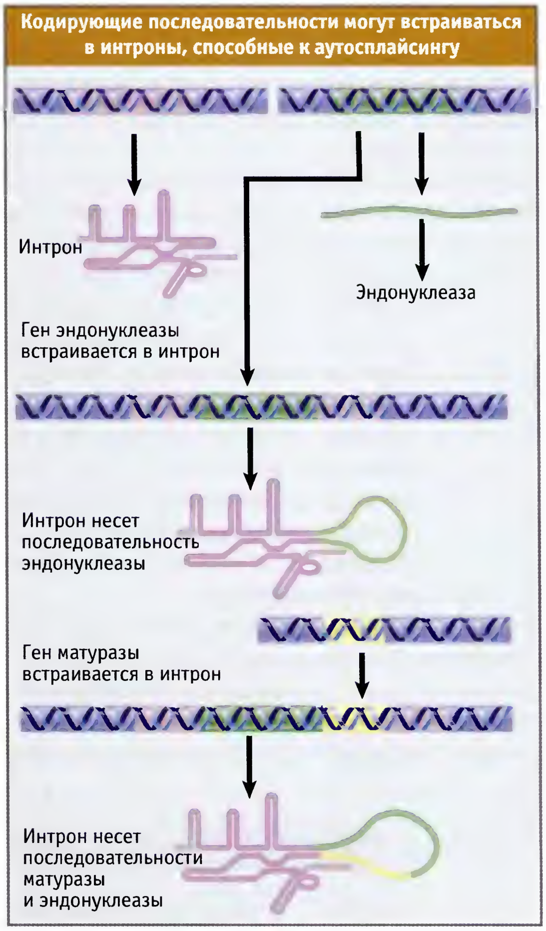


РИС. 27.15 Интрон, кодирующий способную к автосплайсингу РНК, первоначально существует в виде независимой последовательности. В результате встраивания последовательности эндонуклеазы образуется интрон, способный к хомингу и обладающий мобильностью. Последующее встраивание последовательности матуразы усиливает способность интрона сворачиваться в активную структуру, обеспечивающую эффективный сплайсинг

находится в другом домене; он фактически выходит на поверхность белковой глобулы.

Интрон, кодирующий матуразу, не всегда способен вырезать себя из транскрипта без помощи этой матуразы; по сути, она является фактором сплайсинга, специфичным к последовательности, которая его кодирует. Этот фактор содействует правильному фолдингу каталитической сердцевины интрона.

Сочетание в одном белке двух активностей, эндонуклеазы и матуразы, может указывать на путь, пройденный интроном в процессе эволюции. На РИС. 27.15 показано, что прообразом интрона следует считать независимый, автокаталитически сплайсируемый элемент. Позднее в эволюции этому элементу досталась вставка, кодирующая эндонуклеазу. Вставка придала интрону мобильность, но вполне могла нарушить способность РНК-интрона самостоятельно принимать каталитически активную структуру. При этом интрон должен был испытать давление отбора на предмет оптимизации фолдинга

силами специальных белков. Обретение последовательности, кодирующей матуразу, вернуло интрону его первоначальную независимость.

Некоторые интроны группы II не несут в себе информации об активности матуразы, но могут пользоваться белками — продуктами экспрессии других генов — с похожей активностью. Это указывает на возможный путь эволюции факторов основного аппарата сплайсинга. В своем первоначальном виде фактор сплайсинга мог быть матуразой, специфичной к конкретному интрону. Со временем последовательность, кодирующая матуразу, была привнесена за границы своего интрона, потом специфичность фермента распространилась и на другие интроны. Между прочим, малая ядерная РНК (мяРНК; snРНК) вполне может быть одним из продуктов эволюции каталитического ядра интрона.

27.8 Каталитическая активность РНКазы Р обеспечивается РНК

Ключевые понятия:

- Фермент рибонуклеаза Р представляет собой рибонуклеопротеин, каталитическую активность которого следует отнести на счет РНК.

Одной из первых демонстраций разносторонних возможностей РНК послужило раскрытие механизма действия рибонуклеазы Р. Этот фермент принадлежит бактериям *E.coli*; его можно охарактеризовать как эндонуклеазу, осуществляющую процессинг тРНК. Рибонуклеазу Р (или РНКазу Р) можно разделить на два компонента: молекулу РНК размером 375 оснований и полипептид массой 20 кДа. В первых экспериментах по изучению активности этого фермента *in vitro* были выбраны такие условия, при которых для расщепления тРНК оказались необходимыми оба компонента. Впоследствии обнаружилось, что при других условиях (с увеличением концентрации Mg^{2+}) наличие белкового компонента не принципиально. *Катализатором реакции служит именно РНК.* Эту РНК, саму по себе, можно назвать ферментом; каждая молекула «фермента» катализирует расщепление множества молекул субстрата. Несмотря на то что каталитическая активность РНКазы Р совершенно явно заключена в РНК, наличие белкового компонента очень сильно ускоряет ход реакции (см. таблицу на рис. 27.10).

Фермент РНКаза Р является продуктом экспрессии двух генов, кодирующих, соответственно, каталитическую РНК и белковый компонент этого фермента. Мутация в любом из этих генов может инактивировать РНКазу Р в условиях *in vivo*, так что для естественной активности фермент должен содержать оба компонента.

Сперва считалось, что каталитическая активность РНКазы Р обеспечивается белком, а РНК играет какую-нибудь подчиненную роль. Скажем, помогает в связывании субстрата (потому что в ней есть несколько коротких сегментов, комплементарных нескольким открытым сегментам тРНК). Как позднее выяснилось, распределение ролей между белком и РНК в данном случае ровно обратное!

27.9 Вироиды обладают каталитической активностью

Ключевые понятия:

- Вироиды и вирусоиды образуют молоткообразные структуры, способные к авторасщеплению.
- Похожие структуры возникают при взаимодействии цепи-субстрата с комплементарной цепью-ферментом, которая ее разрезает.
- Цепь-фермент, введенная в клетку, может образовать пару с подходящей цепью-мишенью и затем разрезать эту мишень.

Еще один пример способности РНК к проявлению каталитической активности демонстрируют небольшие молекулы, которые можно встретить в клетках растений. Имея размеры ~350 оснований, они квалифицируются как автокаталитические эндонуклеазы. На базе этих молекул можно создавать генно-инженерные конструкции, чтобы продуцировать каталитически активные молекулы РНК, подобные интрону группы I *T. thermophila*, активность которых распространяется на произвольные субстраты.

Небольшие автокаталитические РНК растительного происхождения можно разделить на две группы: вироидов и вирусоидов. **Вироидами** называют инфекционно-способные молекулы РНК, которые не нуждаются в инкапсидации в какую бы то ни было белковую оболочку. **Вирусоиды** похожи на вироиды, но, в отличие от первых, вирусоиды не могут самостоятельно заражать новые клетки — им нужна помощь вирусов-паразитов растений, которые упаковывают их в свои капсиды вместе с копиями собственных геномов. Вирусоиды не способны реплицироваться независимо — им требуется поддержка вирусов; иногда вирусоиды называют **сателлитными РНК**.

Вироиды и вирусоиды реплицируются по типу «катящегося кольца» (см. рис. 16.6). Цепь РНК вирусоида, упакованная в вирусную частицу, расценивается как (+)-цепь. Комплементарная цепь, продукт репликации РНК, — это, соответственно, (-)-цепь. В клетках растений можно обнаружить мультимеры обеих цепей; мономеры (+) и (−) представляют собой нарезку цепей реплицированного материала, который сходит с «катящегося кольца» непрерывной вереницей геномных копий. Кольцевые мономеры (+)-цепи возникают от замыкания (лигирования) концов линейных мономеров.

В условиях *in vitro* (+)- и (−)-цепи вироидов и вирусоидов способны к саморасщеплению. Реакция поддерживается двухвалентными катионами металлов; продукты реакции имеют характерный набор концов: 5′-ОН-конец и конец 2′-3′, замкнутый фосфодиэфирной связью.

Некоторые из рассматриваемых РНК готовы с ходу провести реакцию *in vitro* в физиологических условиях. Другим разновидностям таких РНК требуется предварительное плавление, потому что в обычном виде эти молекулы не годятся для катализа (для принятия нужной конформации им необходимо пройти цикл денатурации и ренатурации).

Чтобы разрезать самого себя, вироид или вирусоид должен образовать вторичную структуру типа «молоток», изображенную в верхней части РИС. 27.16. Структура охватывает не всю молекулу РНК, но лишь окрестность сайта, в который вносится разрез. Для катализа достаточно наличия одного «молотка». Остальная часть молекулы подлежит вырезанию; в этом случае малая РНК легко складывается без предварительной денатурации и спонтанно саморасщепляется. Это говорит о том, что остальные участки молекулы могут тормозить реакцию.

Активный центр РНК вироидов и вирусоидов образован последовательностью из 58 рибонуклеотидов. В «молотке» есть три шпильки; они имеют всегда одни и те же размеры и занимают одни и те же положения. В «молотке» насчитывается 13 консервативных рибонуклеотидов; их положения сосредоточены в центральной части «молотка». Благодаря особому сочетанию консервативных азотистых оснований центральной части с двухцепочечными участками шпилек «молоток» обретает свою каталитическую активность.

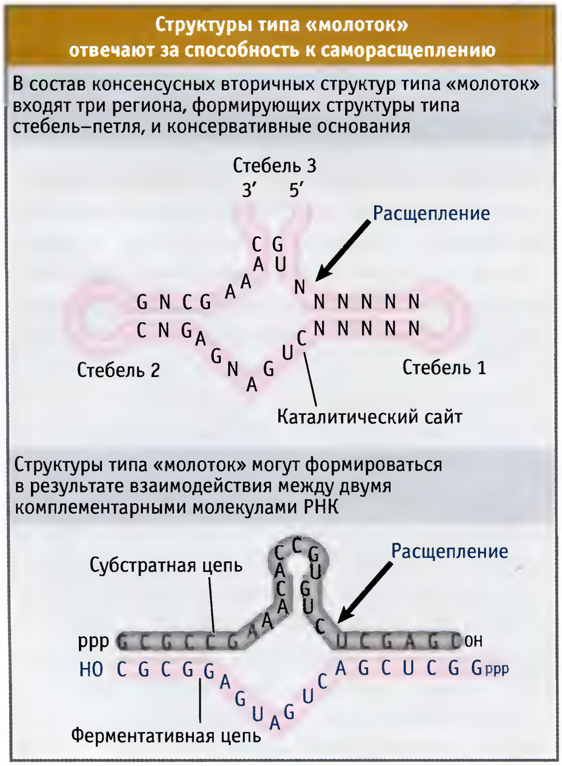


РИС. 27.16 Вироидные и вирусоидные сайты саморасщепления содержат консенсусные последовательности и формируют вторичную структуру типа «молоток» в результате внутримолекулярного спаривания. Такие структуры могут быть также образованы в результате спаривания между субстратной цепью и «ферментативной» цепью

Любопытно, что активный «молоток» также может быть сформирован из двух отдельных молекул РНК: каждая молекула образует свою сторону «молотка». Пример такого «молотка» приведен в нижней части рис. 27.16; в данном случае «молоток» получен от гибридизации 19-членника с 24-членником. Гибрид успешно имитирует структуру «молотка» (несмотря на отсутствие петель I и III на концах соответствующих шпилек). Для того чтобы пошла реакция, достаточно просто добавить 19-членник к 24-членнику; разрез, полученный в результате реакции, будет точно соответствовать своему естественному положению.

В каталитически активном гибриде двух молекул 24-членник, очевидно, играет роль субстрата, в то время как 19-членник играет роль, собственно, фермента. Смешивание небольшого количества 19-членника (фермента) с избытком 24-членника (субстрата) приводит к разрезанию всех молекул 24-членника. Это говорит о том, что взаимодействие субстратов с ферментом носит циклический характер; каждый цикл состоит из формирования «молотка», внесения разреза и диссоциации 19-членника с фрагментами разрезанного 24-членника. Совершенно очевидно, что в ходе реакции фермент не расходуется, т. е. здесь мы имеем дело с истинным катализом. К тому же по всем своим характеристикам эта реакция похожа на другие реакции, где катализатором выступает РНК (см. таблицу на рис. 27.10). Следовательно, 19-членник с полным правом может быть причислен к рибозимам с эндонуклеазной активностью.

Анализ кристаллической структуры показал наличие у «молотка» V-образной пространственной структуры. Активный центр приходится на крутой изгиб в структуре «молотка»; это схематично изображено на РИС. 27.17. Ион Mg^{2+}, находящийся в активном центре, играет исключительно важную роль: он связан с цитидином мишени и цитидином у основания шпильки I (и, возможно, еще с соседним уридином). Ион Mg^{2+} вытесняет протон из 2′-ОН цитидина мишени и затем открыто атакует лабильную фосфодиэфирную связь. Мутации в рибонуклеотидной последовательности молотка, специфически нарушающие реакцию в ее переходной стадии, могут приходиться и на активный центр, и на другие части «молотка». Это говорит о том, что, возможно, самой реакции всегда предшествуют какие-то изменения во вторичной и третичной структуре «молотка».

Эндонуклеолитическое расщепление РНК в исполнении другой РНК, рибозима, очень похоже на рестрикцию (своей специфичностью; только мишенью в данном случае служит не ДНК, а РНК). В теории можно придумать множество различных сочетаний из двух молекул РНК, которые будут образовывать друг с другом «молоток», так что одна из них будет играть роль фермента, а другая — роль субстрата. Для многих сочетаний такое распределение ролей подтверждено практикой. Если ввести такой фермент в живую клетку, реакцию можно

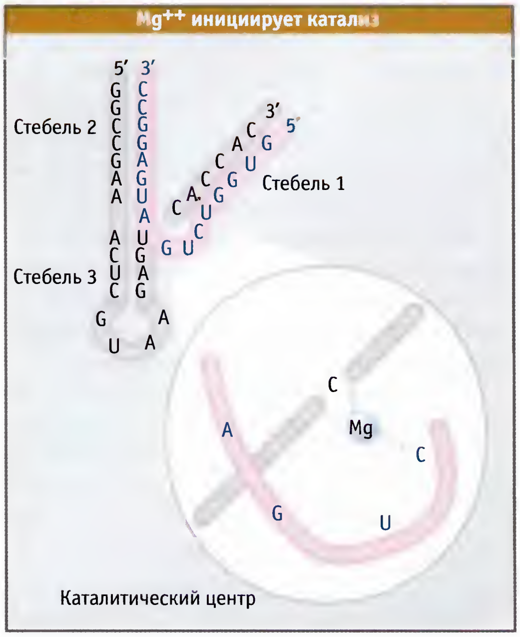


РИС. 27.17 Рибозим типа «молоток» формирует V-образную пространственную структуру, в которой стебель 2 располагается над стеблем 3. Каталитический центр расположен между стеблями 2, 3 и 1. Он содержит ион магния, инициирующий реакцию гидролиза

детектировать *in vivo* (при условии, что в клетке есть подходящая мишень). С точки зрения разработки новых методов исследования, это открывает широкий спектр возможностей, поскольку рибозим, помещенный под контроль регулируемого промотора, можно использовать, подобно антисмысловым РНК, для прицельного выключения экспрессии избранных генов в заданных условиях.

27.10 Редактирование РНК касается конкретных азотистых оснований

Ключевые понятия:

- **В незрелых мРНК аполипопротеина В и рецепторов глутамата происходит сайт-специфичное дезаминирование; катализируемое цитидин- и аденозин-дезаминазами; реакции приводят к изменению смысла кодирующей области.**

Долгое время считалось, что последовательность мРНК всегда является точным отображением последовательности геномной ДНК. Центральная догма молекулярной биологии подразумевает точную однонаправленную передачу генетической информации между компонентами аппарата экспрессии: непрерывная нуклеотидная последовательность ДНК транскрибируется в непрерывную последовательность мРНК, которая, в свою очередь, транслируется в белок. Открытие прерывистых генов (и, как следствие, потребность в специальном механизме удаления интронов) добавило в самую середину этой информационной лестницы еще одну ступень. Оказалось, что в очень многих случаях разрозненные кодирующие последовательности, взятые прямо из генома, не могут транслироваться в белок до тех пор, пока не будут состыкованы. Тем не менее процесс стыковки (сплайсинг РНК) никак не меняет интерпретацию генетической информации, записанной в ДНК; он просто доводит генетическую информацию до состояния, поддающегося интерпретации (т. е. трансляции в белок).

Запланированные изменения в информационном содержании геномных последовательностей происходят крайне редко, в исключительных обстоятельствах. Самым ярким примером служит возникновение в лимфоцитах зверей и птиц новых геномных последовательностей, кодирующих иммуноглобулины (об этом можно прочитать в гл. 23 «Как формируется многообразие антител»). Как результат процесса реконструкции гена, кодирующего иммуноглобулин, в геномной ДНК отдельно взятой клетки возникает новая информация. *Аминокислотная последовательность конкретного полипептида изменяется в результате соматической мутации.* Тем не менее все транскрипты иммуноглобулинов являются абсолютно достоверными матричными копиями геномной ДНК; ни о каком редактировании этих копий здесь речи не идет.

Редактирование РНК — это *изменение в информационном содержании кодирующих последовательностей на уровне транскриптов (мРНК или молекул-предшественников мРНК).* Дело в том, что в некоторых случаях кодирующая последовательность РНК отличается от своего эквивалента в геномной ДНК. Насколько известно, в ходе эволюции подобные ситуации складывались дважды. Во-первых, в некоторых клетках млекопитающих, в транскриптах отдельно взятых генов, происходят обязательные точечные замены конкретных азотистых оснований РНК (1–2 замены на транскрипт), приводящие к изменению в аминокислотной последовательности кодируемого белка. Во-вторых, в митохондриях паразитических простейших *Trypanosoma* с транскриптами происходят очень обстоятельные изменения, касающиеся обширных участков в транскриптах нескольких генов.

На РИС. 27.18 даны схемы последовательностей, кодирующих белок аполипопротеин В (apoB) в клетках двух различных тканей мыши (кишечника и печени). Для apoB в геноме мыши есть один-единственный ген, во всех тканях один и тот же; его кодирующая область состоит из 4563 кодонов. Транскрипция этого гена дает мРНК, во всех тканях она одного размера. В клетках печени она транслируется в белок массой 512 кДа; аминокислотная последовательность этого белка соответствует полноразмерной кодирующей области гена.

В клетках кишечника дело обстоит иначе. Там синтезируется укороченная версия apoB, массой всего ~250 кДа; она соответствует N-концевой части полноразмерного

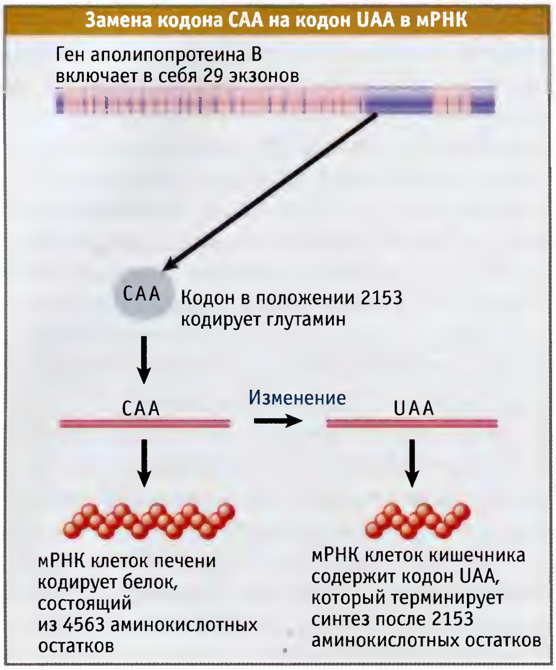


РИС. 27.18 Клетки кишечника и клетки печени характеризуются одинаковой нуклеотидной последовательностью гена *apoB*. Однако в результате замены основания и образования терминирующего кодона мРНК в клетках кишечника отличается от мРНК в клетках печени

белка. Укорачивание продукта трансляции мРНК apoB является результатом редактирования, специфичного для клеток кишечника. В положении 2153 мРНК *apoB* происходит замена C на U; в результате кодон CAA, соответствующий глютамину, превращается в стоп-кодон охра.

Отчего происходит эта замена? (Несомненно, она происходит именно на уровне РНК, поскольку в геноме нет никаких альтернативных генов или экзонов, и паттерн сплайсинга во всех клетках тоже одинаковый.) Приходится признать, что замена вносится непосредственно в транскрипт на одной из стадий процессинга.

Еще один похожий случай связан с экспрессией рецепторов глютамата в мозге крысы. Похожим образом, редактирование по одному из положений РНК меняет один из кодонов, что ведет к включению в полипептид вместо одного из глутаминов аргинина. С точки зрения функции, замена аминокислоты влияет на проводимость ионного канала (присовокупленного к собственно рецепторной части этого белка) и, таким образом, сильно влияет на способность данного рецептора к участию в передаче нервного импульса. В другом положении того же полипептида аргинин может быть заменен на глицин, опять же, посредством редактирования транскриптов на одной из стадий процессинга.

Как выглядят реакции редактирования РНК? В транскриптах *apoB* идет замена C_{2153} на U; в транскриптах рецептора глутамата обе поправки представляют собой замены A на I (инозин). Каждая замена — это результат реакции *дезаминирование*, в которой азотистое основание нуклеотида теряет аминогруппу. Катализаторами дезаминирования выступают специальные ферменты, цитидин- и аденозин-дезаминазы, соответственно. Этот тип редактирования транскриптов особенно характерен для различных отделов нервной системы млекопитающих. (Интересно, что у *Drosophila melanogaster* тоже есть потенциальные мишени цитозин-дезаминазы; их известно 16, и все они находятся в генах, задействованных в передаче нервного импульса.) Во многих случаях редактирование обеспечивает замену аминокислоты в положении чрезвычайной функциональной значимости.

Чем обусловлена высокая специфичность редактирования? Специфичности ферментов, осуществляющих дезаминирование, здесь явно не достаточно. (Например, самая изученная из дезаминаз, аденозин-дезаминаза, будет работать с любым основанием A при условии, что оно содержится в двухцепочечном участке вторичной структуры РНК.) Поэтому ферменты, которые осуществляют редактирование, хоть и похожи на дезаминазы общего профиля, имеют специальные домены (или дополнительные субъединицы), усиливающие их специфичность. Например, у комплекса, редактирующего РНК apoB, каталитическая субъединица похожа на бактериальную цитидин-дезаминазу, но, в отличие от нее, эта субъединица имеет дополнительный РНК-связывающий мотив, который помогает всему комплексу узнавать конкретную мишень редактирования. То же самое в нервной ткани: для каждой мишени в транскриптах рецепторов глутамата и серотонина существует своя особая разновидность аденозин-дезаминазы, работающая с высокой специфичностью.

Какова природа специфичности? Некоторые комплексы редактирования могут распознавать определенные регионы во вторичной структуре РНК (этим они похожи на ферменты, модифицирующие тРНК). Другие комплексы редактирования распознают конкретную нуклеотидную последовательность. Так, эксперименты показали, что для грамотного редактирования РНК по типу apoB в условиях *in vitro* достаточно наличия у сай-

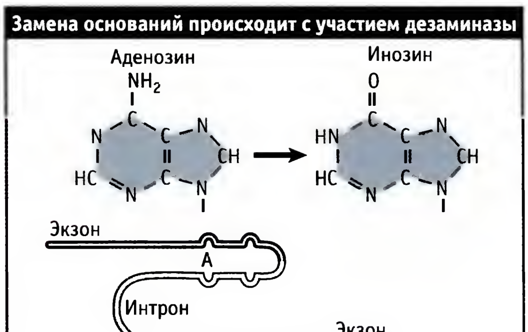


РИС. 27.19 Изменение мРНК происходит при действии деаминазы на аденин, расположенный в участке, где РНК образует неполностью спаренный дуплекс

та-мишени РНК небольшого сегмента (размером ~26 оснований). На РИС. 27.19 изображена другая ситуация: там показано, что для успешного редактирования РНК рецептора GluR-B должно быть выполнено условие взаимодействия редактируемого сегмента с комплементарным сегментом нижележащего интрона. В данном случае для специфичного узнавания комплексом редактирования своей мишени в РНК очень важен паттерн неспаренных оснований внутри двухцепочечного участка, образующегося от взаимодействия двух комплементарных сегментов. Таким образом, каждый комплекс редактирования предъявляет свои требования к потенциальным субстратам; жесткость этих требований служит залогом высокой специфичности точечного редактирования транскриптов.

27.11 Редактирование транскриптов направляется специальными малыми РНК

Ключевые понятия:

- В митохондриях трипаносом имеет место экстенсивное редактирование РНК: транскрипты конкретных генов претерпевают многочисленные вставки и делеции уридинов.
- Две молекулы, транскрипт и наводящая РНК, взаимодействуют своими комплементарными участками; эти участки располагаются по обе стороны от последовательности-мишени редактирования.
- Матрицей для добавления (реже, удаления) уридинов служит наводящая РНК.
- Катализ редактирования осуществляется белковым комплексом, который обладает активностями эндонуклеазы, терминальной уридил-трансферазы и РНК-лигазы.

Совершенно иной тип редактирования РНК имеет место в митохондриях простейшего *Trypanosoma*. Первый случай, который был открыт, это аминокислотная последовательность субъединицы II цитохромоксидазы, которая соответствует кодирующей области гена *coxII*, в середине которой произошел сдвиг открытой рамки считывания. Последовательности гена и белка (нуклеотидная и аминокислотная соответственно) приведены на РИС. 27.20. В данном случае, у разных видов трипаносом, происходит явное отступление от традиционного алгоритма интерпретации генетической информации. Какова же природа этого изменения? Как работает этот ген?

По сравнению со своей геномной матрицей, зрелая мРНК *coxII* содержит четыре дополнительных рибонуклеотида. Это уридины, вставленные в кодирующую последовательность по месту сдвига открытой рамки считывания. (Поскольку никакого другого гена с такой последовательностью в митохондриальном геноме нет, приходится признать, что дополнительные рибонуклеотидные положения возникают во время транскрипции либо после нее.)

Такой тип редактирования РНК был зафиксирован для целого ряда митохондриальных генов *Trypanosoma*; зачастую редактирование заключается не только в добавлении, но и в удалении уридинов. РИС. 27.21 посвящен экстремальному случаю редактирования транскриптов гена *coxIII Trypanosoma brucei*.

Как ни удивительно, мРНК *coxIII* больше чем наполовину состоит из уридинов, которых нет в гене. Сравнение мРНК с ее геномной матрицей показало, что у них *нет ни одного общего сегмента протяженностью более семи нуклеотидов*! Зато в мРНК *coxIII* насчитывается множество вставок из одних уридинов; число уридинов в этих вставках доходит до семи. У парамиксовирусов, SV5 и вируса кори, обнаружены аналогичные расхождения последовательности мРНК со своей ДНК-матрицей (только в транскрипт добавляются не U, а G).

Каков механизм возникновения специфичных вставок уридинов? Несмотря на кажущуюся беспорядочность, для каждого гена паттерн вставок и делеций всегда один и тот же. Для каждого редактируемого гена существует специальная молекула, так называемая **наводящая РНК**. Она содержит последователь-

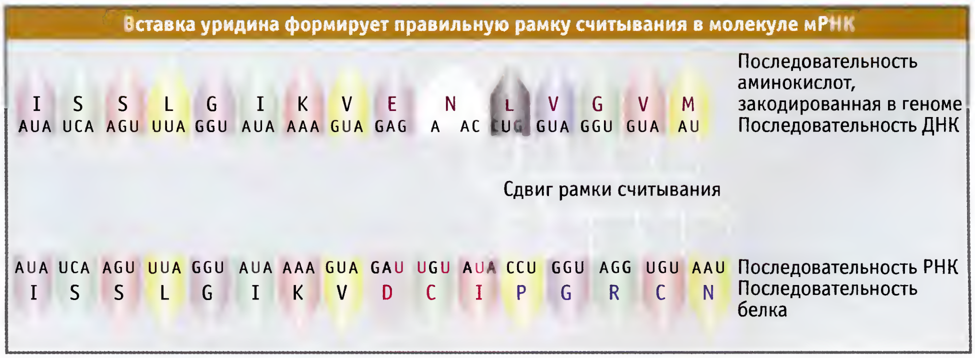


РИС. 27.20 мРНК гена *coxII* трипаносомы характеризуется сдвигом рамки считывания относительно ДНК; корректная рамка считывания формируется в результате вставки четырех уридинов

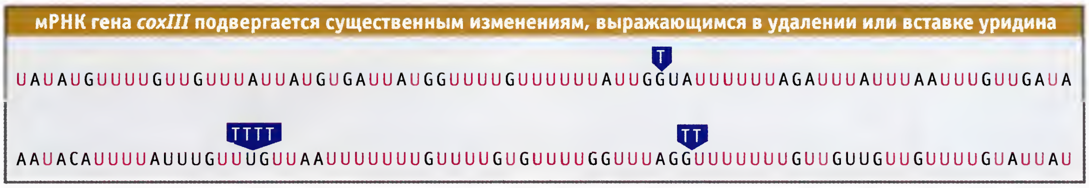


РИС. 27.21 Частичная последовательность мРНК гена *coxIII T. brucei* показывает наличие большого количества уридинов, не кодируемых ДНК (отмечены красным цветом) или, напротив, удаленных из последовательности РНК (отмечены как Т)

ность, комплементарную окончательной (отредактированной) версии мРНК. На РИС. 27.22 представлена модель работы наводящей РНК для гена цитохрома *b* другого паразитического простейшего (из рода *Leishmania*). Последовательность в верхней части рисунка соответствует оригинальной версии транскрипта (до редактирования). Пробелы показывают, где именно, в процессе редактирования, будут появляться вставки. Для получения валидной РНК в эту область транскрипта должно быть вставлено, в общей сложности, восемь уридинов.

Наводящая РНК и транскрипт имеют довольно протяженный участок комплементарности; этот участок перекрывает мишень редактирования. Как правило, комплементарность распространяется в 3'-концевом направлении от мишени редактирования дальше, чем в 5'-концевом. Взаимодействие наводящей РНК с нередактированным транскриптом оставляет пробелы в местах, где избыточные А в наводящей РНК не находят себе комплементарных партнеров. Наводящая РНК служит матрицей; в соответствии с ней транскрипт получает вставки уридинов (комплементарных избыточным аденозинам наводящей РНК). После завершения всех реакций редактирования наводящая РНК отделяется от мРНК, которая теперь годится для трансляции.

Редактирование (спецификация окончательной версии мРНК) может протекать довольно сложно. В приведенном примере в протяженный участок транскрипта добавляется, в общей сложности, 39 уридинов. Такое фундаментальное редактирование выполняется при участии *двух* наводящих РНК, которые перекрывают друг друга. Сначала одна из них занимает 3'-концевую

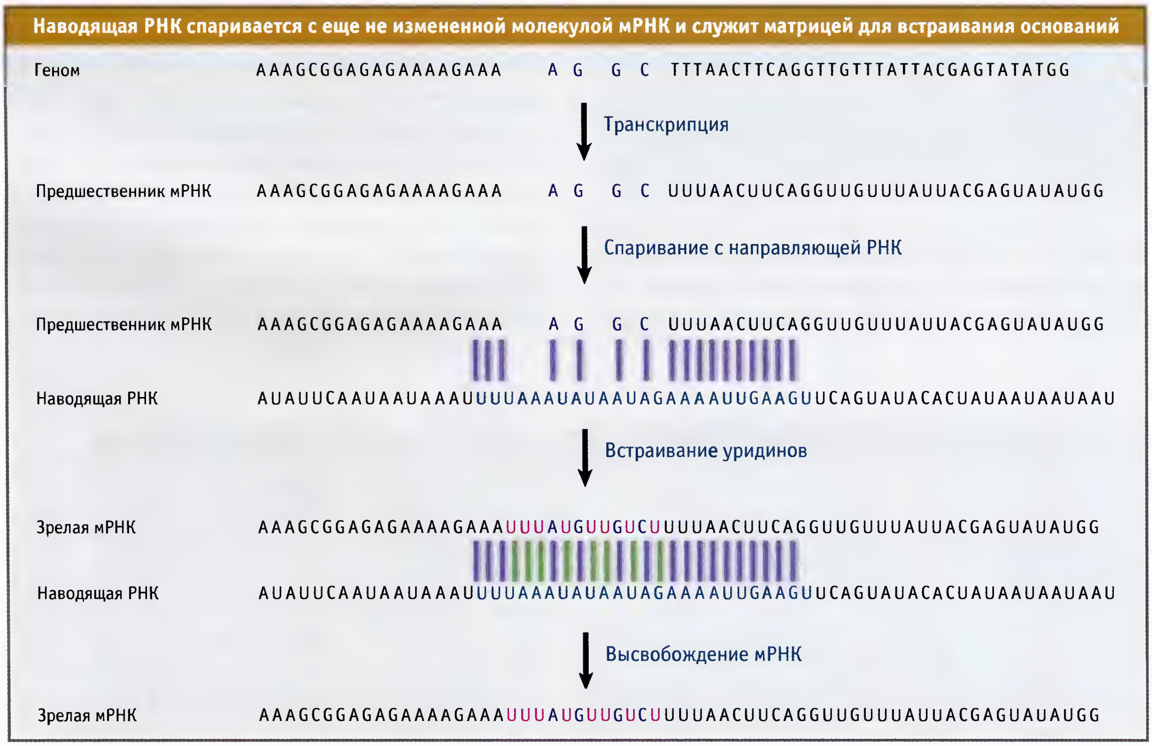


РИС. 27.22 Еще не подвергшаяся изменениям РНК спаривается с наводящей молекулой РНК в участках, ограничивающих изменяемый регион. Наводящая РНК служит матрицей для встраивания уридинов. мРНК, образующаяся в результате встраивания оснований, комплементарна наводящей РНК

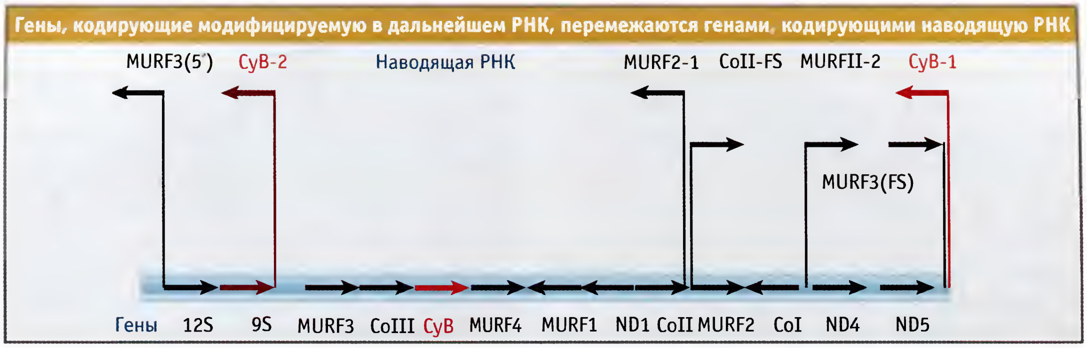


Рис. 27.23 Геном *Leishmania* содержит гены, кодирующие модифицируемую в дальнейшем РНК. Эти гены перемежаются генами, кодирующими наводящую РНК, необходимую для образования мРНК корректной последовательности. Созревание мРНК некоторых генов требует участия нескольких наводящих РНК. В модификации мРНК гена *CyB*, кодирующего цитохром b, участвуют наводящие РНК, кодируемые двумя генами – *CyB-1* и *CyB-2*

часть участка и, сделав свое дело, уходит, предоставляя второй наводящей РНК завершить редактирование оставшейся части.

Откуда же берутся наводящие РНК? Они представляют собой продукты специальных, независимых транскрипционных единиц. На рис. 27.23 дана карта соответствующей области мтДНК *Leishmania*. Эта область содержит «ген» цитохрома *b* (кавычки напоминают, что последовательность гена нуждается в серьезной доработке за счет редактирования РНК). Та же область содержит еще две транскрипционных единицы, продуктами которых являются наводящие РНК. Из рисунка видно: матрицы для синтеза наводящих РНК перемежаются кодирующей областью самого «гена».

Мутации, от которых смысл мРНК (т.е. аминокислотная последовательность продукта трансляции) оказывается измененным за счет нарушений редактирования, могут находиться как в самом «гене», так и в наводящих РНК. По всем правилам классической генетики, каждую из этих матриц следует считать составной частью «гена». Экспрессия двух транскрипционных единиц для наводящих РНК происходит независимо; как следствие, в *транс*-сочетаниях они должны дополнять друг друга. Анализируя всю совокупность мутаций, можно установить наличие трех групп комплементации; каждая группа мутаций маркирует одну из транскрипционных единиц, необходимых для правильного кодирования аминокислотной последовательности одного полипептида.

Характерные особенности частично отредактированных полупродуктов процессинга говорят о том, что процесс редактирования охватывает не всю молекулу РНК сразу, но постепенно, в направлении 3′–5′. Взаимодействуя с интактным, нередактированным, транскриптом, наводящие РНК задают паттерн уридиновых вставок. Образование этих вставок катализируется специализированным комплексом, с константой седиментации 20S. Комплекс обладает активностями эндонуклеазы, терминальной уридилтрансферазы (или ТУТазы), а также активностью РНК-лигазы. Все эти активности изображены на рис. 27.24. Комплекс занимается рекрутингом наводящих РНК, то есть привлекает их к связыванию еще не редактированного транскрипта, который служит комплексу субстратом. Чтобы добавить в транскрипт (или удалить из него) уридиновый остаток, комплекс должен разрезать транскрипт в том месте, где он *не* взаимодействует с наводящей РНК. После внесения разреза происходит вставка (или делеция) уридина. *Вставка (делеция) сообщает транскрипту комплементарность к наводящей РНК в данном конкретном месте.* Последующее лигирование восстанавливает целостность субстрата. Затем весь цикл реакций редактирования повторяется — для следующего по порядку сайта.



Рис. 27.24 Модификация мРНК является результатом следующих процессов: расщепление РНК, удаление или добавление остатков U, лигирование концов. Реакция катализируется комплексом ферментов при участии направляющей РНК

Источником добавляемых уридинов служат молекулы уридинтрифосфатов (УТФ). Само добавление обеспечивается активностью ТУТазы. (Возникновение делеций пока не выяснено, возможно, они продуцируются ТУТазой или какой-то особой экзонуклеазой.) Одно время считалось, что цепочка (поли)U, присутствующая на конце наводящей РНК, может служить для редактирования: либо как источник дополнительных U, либо для накопление избытка U. В дальнейшем эта гипотеза не подтвердилась; образование «хвоста» (поли)U есть не что иное, как аберрантная реакция, не имеющая прямого отношения к редактированию.

Судя по типу структуры, свойственной частично отредактированным молекулам, уридины совершенно точно добавляются по одному, а не группами. Процесс редактирования выглядит как последовательность рабочих циклов. В каждом цикле происходит добавление нуклеотида U; затем происходит его проверка на комплементарность матрице (наводящей РНК). Цикл завершается удалением либо сохранением добавленного U (по результатам проверки). Согла̂сно этой модели, отредактированная версия транскрипта возникает *постепенно*.

Редактирование с добавлением нуклеотидов G, свойственное вирусам, скорее всего, происходит по сходным механизмам, но это пока не установлено.

27.12 Сплайсинг белка — автокаталитическая реакция

Ключевые понятия:

- Участок полипептида, обозначаемый как интеин, может катализировать свой собственный выход из состава белка; при этом фланкирующие участки, экстеины, смыкаются друг с другом; это событие известно как сплайсинг белка.
- Катализатором сплайсинга белка служит интеин.
- У большинства интеинов есть две независимых активности: кроме участия в сплайсинге белка, многие интеины проявляют себя как хоминг-эндонуклеазы.

По своему результату **сплайсинг белка** полностью аналогичен сплайсингу РНК. Аминокислотная последовательность только что синтезированного полипептида является отображением открытой рамки считывания зрелой мРНК. Созревая до функционального белка, полипептид может утрачивать отдельные участки этой последовательности. Удаление этих участков и есть сплайсинг.

Удаляемые участки полипептида называют **интеинами** (inteins; от слов «**in**tron» и «pro**tein**»). Остальные участки полипептида составляют зрелый белок; их называют **экстеинами** (exteins; от слов «**ex**on» и "pro**tein**"). Аппарат сплайсинга белка не имеет никакого отношения к аппарату сплайсинга РНК. На РИС. 27.25 показан первичный продукт трансляции мРНК, полипептид-предшественник, содержащий интеин, подлежащий вырезанию. На сегодняшний день для белков известно около сотни примеров сплайсинга, у самых разных организмов. Типичный продукт трансляции, который подлежит сплайсингу, содержит один-единственный интеин.

Самый первый из известных интеинов был обнаружен в полипептиде-предшественнике ДНК-полимеразы археబактерий. В транскрибируемой части гена, кодирующего этот фермент, обнаружилась последовательность, никак не отображенная в белке, но и не похожая на интрон. Расследование этого странного обстоятельства показало, что очищенный полипептид, первичный продукт трансляции, *способен вырезать соответствующий участок* в автокаталитической манере. Хотя интеин ведет катализ реакции вполне автономно, ее эффективность очень сильно зависит от экстеинов. Реакция не требует энергозатрат и представляет собой набор перестроек в паттерне химических связей; все это показано на РИС. 27.26.

В первой реакции белкового сплайсинга OH- или SH-группа в радикале первой аминокислоты интеина атакует пептидную связь, соединяющую эту аминокислоту с последней аминокислотой первого экстеина. В результате атаки происходит перенос экстеина с концевой аминогруппы интеина на радикал. Образовавшийся стык выглядит как ацильная связь N-O или N-S. Эта связь подлежит атаке со стороны OH- либо SH- группы на радикале первой аминокислоты теперь уже второго экстеина, которая приводит к тому, что первый экстеин переносится на радикал N-концевой аминокислоты второго экстеина. После этого C-концевой аспарагин интеина циклизуется, и концевая NH-группа второго экстеина атакует ацильную связь для замены ее на стандартную пептидную связь. Каждая из перечисленных реакций мо-



РИС. 27.25 Белковый сплайсинг обеспечивает удаление из последовательности белка интеинов и соединение экстеинов между собой

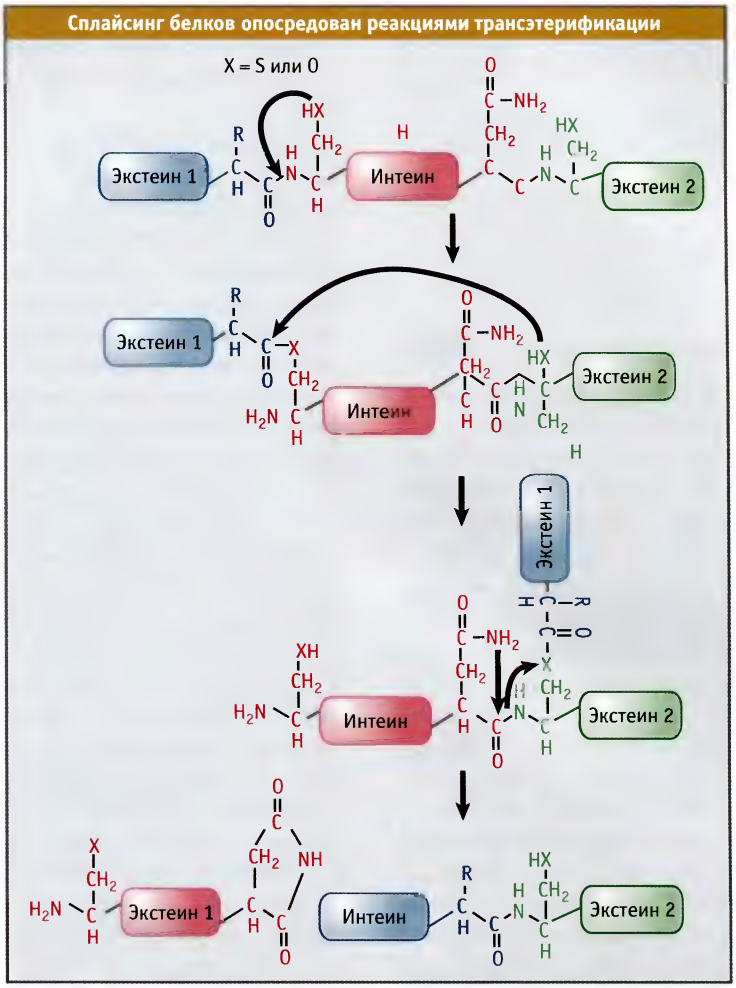


РИС. 27.26 **Образование новых связей происходит в результате серии реакций трансэтерификации, в которых задействованы OH-группы серина или треонина или SH-группа цистеина. В ходе сплайсинга экстеины соединяются между собой пептидной связью, а интеин, у которого формируется кольцевой C-конец, высвобождается**

жет протекать самопроизвольно, но в таком случае она характеризуется очень низкой скоростью. Для того чтобы реакции протекали координированно и достаточно быстро, необходим катализатор — интеин.

У интеинов есть ряд характерных особенностей. В кодирующей области гена они выглядят как вставки, не отображаемые в аминокислотную последовательность полипептида, которые при этом не сбивают рамку трансляции. Наличие у гена хотя бы одного гомолога, не имеющего такой вставки, облегчает идентификацию интеинов. В любом интеине N-концевой аминокислотный остаток — это обязательно либо серин, либо цистеин (поскольку радикалы серина и цистеина обладают «реактивными» группами –XH). C-концевой остаток интеина — это обязательно аспарагин. Каталитическая активность типового интеина обеспечивается его N-концевой частью (размером ~150 аминокислотных остатков) и C-концевой частью (~50 остатков). Середина интеина может иметь свой собственный набор функций.

У многих интеинов есть одно удивительное свойство. Они обладают эндонуклеазной активностью, специфичной к определенной последовательности-мишени в ДНК. Выступая в роли хоминг-эндонуклеазы, интеин узнает в ДНК свою мишень и надрезает ее для вставки. (Функции хоминг-эндонуклеаз см. на рис. 27.12 в разд. 27.5 «Некоторые интроны группы I кодируют эндонуклеазы, поддерживающие их мобильность».) Участие в сплайсинге белка и внесение надрезов в конкретные мишени ДНК — это две независимых активности интеина.

Наличие у интеина двух столь разных активностей на самом деле неожиданно и (на сегодняшний день) не очень хорошо объяснимо. На этот счет были высказаны две гипотезы. Одна из них гласит, что исходно две активности были взаимосвязаны, а затем стали независимыми; при этом некоторые интеины потеряли свою эндонуклеазную активность. Вторая гипотеза делает ставку на инвазию эндонуклеазной активности в интеи-

ны, возникшие как функциональные единицы сплайсинга белков. Причина инвазии пока не нашла объяснения; тем не менее вторая гипотеза хорошо согласуется с установленным фактом инвазии хоминг-эндонуклеаз в другие типы каталитически активных единиц (например, в интроны первой группы).

27.13 Резюме

Молекулы РНК, способные к катализу биохимических реакций, называют рибозимами. Автокаталитический сплайсинг РНК (автосплайсинг) свойственен двум группам интронов, широко представленным у низших эукариот, а также у прокариот и в митохондриях. Все структурное обеспечение, необходимое для катализа, содержится в самом интроне. Таким образом, вся информация, необходимая для реакций сплайсинга, заключена в рибонуклеотидной последовательности интрона. Однако в условиях *in vivo* эти реакции не идут без поддержки специальных белков.

Интроны двух групп, I и II, становятся эффективными катализаторами за счет образования особой вторичной и третичной структуры РНК с большим количеством коротких консервативных двухцепочечных участков. Интроны первой группы образуют в себе подобие активного центра, удерживающего РНК субстрат за счет взаимодействий по принципу комплементарности. Элемент рибонуклеотидной последовательности интрона, который занимается этим удержанием, обозначается как ВНП; другие консервативные элементы интрона, в сочетании друг с другом, образуют второй компонент активного центра, где присутствует сайт связывания гуанозина (рибонуклеотида G). Сплайсинг начинается реакцией трансэтерификации, в которой участвует кофактор — гуанозиннуклеотид. Эта реакция не требует притока энергии. Кофактор гуанозиннуклеотид разрушает фосфодиэфирную связь на стыке экзона с интроном; после этого он остается прикрепленным к 5′-концу интрона. Во второй реакции сплайсинга свободный гидроксил, появившийся теперь на 3′-конце первого экзона, атакует стык интрона со вторым экзоном и превращает его в стык между экзонами. Интрон уходит и замыкается в кольцо, теряя концевой участок (~15 рибонуклеотидов вместе с гуанозином, бывшим кофактором). Ряд схожих реакций запускается атакой концевого G-OH-остатка интрона на внутренние фосфодиэфирные связи. В лабораторных условиях этот механизм используется для создания множества различных рибозимов. В условиях *in vitro* такие рибозимы могут выступать катализаторами широкого спектра реакций (например, рибозим может выступать в роли нуклеотидил-трансферазы).

В митохондриях некоторые интроны групп I и II содержат открытые рамки считывания. Белки, кодируемые интронами первой группы, обладают активностью эндонуклеаз: они вносят надрезы по обеим цепям сайта-мишени в ДНК. Внесение надрезов инициирует конверсию гена с копированием интрона в виде вставки в сайт-мишень. Белки, кодируемые интронами второй группы, сочетают в себе несколько активностей. Это эндонуклеазная активность (которая инициирует процесс транспозиции), а также активность обратной транскриптазы (обеспечивающая копирование интрона с РНК в геномный сайт-мишень). Поэтому считается, что большинство интронов групп I и II возникли как вставки. Белки, кодируемые интронами обеих групп, могут обладать активностью матуразы; эта активность очень важна для сплайсинга. Матураза стабилизирует вторичную и третичную структуры каталитически активного интрона РНК; тем самым матураза обеспечивает устойчивую конфигурацию активных центров рибозима.

Способность к катализу биохимических реакций есть у многих РНК. Фермент РНКаза Р представляет собой рибонуклеопротеин; его каталитическая активность обеспечивается РНК-компонентом, тогда как белковый компонент играет лишь вспомогательную роль. Таким образом, РНКазу Р следует считать, скорее, рибозимом, чем классическим ферментом белковой природы. Способность к автокаталитическому расщеплению свойственна многим вирусным РНК; они образуют особую вторичную структуру, которая по очертаниям напоминает молоток. Такие «молотки» могут образовываться не только одной, но и двумя молекулами РНК; в этом случае имеет место разделение ролей: одна из молекул играет роль субстрата, другая — рибозима. Расщепление субстрата идет очень специфично к его последовательности. Специфичность (как точность наводки на конкретную мишень — нуклеотидную последовательность) является характерной особенностью функции любого рибозима.

Редактирование РНК представляет собой особую форму процессинга транскриптов некоторых генов. У млекопитающих и, возможно, у некоторых других животных редактирование идет как замена отдельных азотистых оснований в конкретных положениях. Реакции дезаминирования переводят C в U, A в I. Каждую такую реакцию контролирует специальный комплекс; его каталитическая субъединица похожа на обычную цитидин- или аденозиндезаминазу, однако весь комплекс обладает очень узкой специфичностью к конкретной последовательности-мишени.

В митохондриях трипаносом, а также у парамиксовирусов, имеет место другой тип редактирования, с добавлением (реже, удалением) множества рибонуклеотидов (чаще всего, уридинов). Особенно экстенсивное редактирование наблюдается в митохондриях трипаносом, где некоторые зрелые мРНК наполовину состоят из рибонуклеотидов, встроенных не при транскрипции, а уже в процессе редактирования транскрипта. В данном случае для редактирования используется специальная матрица — так называемая наводящая РНК, края которой комплементарны краям редактируемой области транскрипта. Реакцию ведет целый комплекс из ферментов с набором активностей эндонуклеазы, терминальной

уридил-трансферазы, РНК-лигазы. Комплекс использует свободные рибонуклеотиды U в качестве материала для вставок (возможно, тот же комплекс выбрасывает свободные U после удалений).

Сплайсинг белка — это особое явление, которое не имеет прямого отношения к каталитическим РНК. Тем не менее, мы рассматриваем сплайсинг белка в этой главе, поскольку это автокаталитический процесс, аналогичный сплайсингу РНК — по смыслу, по общему плану молекулярных перестроек, а также по эволюционным параллелям. Подобно сплайсингу РНК, сплайсинг белка не требует притока энергии. Интеин — это участок аминокислотной последовательности, вырезающей сам себя из окружения фланкирующих участков — экстеинов. Многим интеинам свойственна активность хоминг-эндонуклеазы, специфичной к конкретному сайту геномной ДНК; ни к какому сплайсингу эта активность интеина прямого отношения не имеет.

Литература

27.2 Интроны группы I самоустраняются по механизму трансэтерификации

Обзор

Cech, T. R. (1985). Self-splicing PHK: implications for evolution. *Int. Rev. Cytol.* 93, 3–22.

Cech, T. R. (1987). The chemistry of self-splicing PHK and PHK enzymes. *Science* 236, 1532–1539.

Исследование

Been, M. D. and Cech, T. R. (1986). One binding site determines sequence specificity of *Tetrahymena* pre-rPHK self-splicing, *trans*-splicing, and PHK enzyme activity. *Cell* 47, 207–216.

Belfort, M., Pedersen-Lane, J., West, D., Ehrenman, K., Maley, G., Chu, F., and Maley, F. (1985). Processing of the intron-containing thymidylate synthase (td) gene of phage T4 is at the PHK level. *Cell* 41, 375–382.

Cech, T. R. et al. (1981). *In vitro* splicing of the rPHK precursor of *Tetrahymena*: involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell* 27, 487–496.

Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., and Cech, T. R. (1982). Self-splicing PHK: autoexcision and autocyclization of the ribosomal PHK intervening sequence of *Tetrahymena*. *Cell* 31, 147–157.

Myers, C. A., Kuhla, B., Cusack, S., and Lambowitz, A. M. (2002). tPHK-like recognition of group I introns by a tyrosyl-tPHK synthetase. *Proc. Natl. Acad. Sci. USA* 99, 2630–2635.

27.3 Интроны группы I имеют характерную вторичную структуру

Исследование

Burke, J. M. et al. (1986). Role of conserved sequence elements 9L and 2 in self-splicing of the *Tetrahymena* ribosomal PHK precursor. *Cell* 45, 167–176.

Michel, F. and Wetshof, E. (1990). Modeling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J. Mol. Biol.* 216, 585–610.

27.4 Рибозимы проявляют различные каталитические активности

Обзор

Cech, T. R. (1990). Self-splicing of group I introns. *Annu. Rev. Biochem.* 59, 543–568.

Исследование

Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A., and Breaker, R. R. (2004). Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* 428, 281–286.

27.5 Некоторые интроны группы I кодируют эндонуклеазы, поддерживающие их мобильность

Обзор

Belfort, M. and Roberts, R. J. (1997). Homing endonucleases: keeping the house in order. *Nucleic Acids Res.* 25, 3379–3388.

27.6 Интроны группы II могут кодировать многофункциональные белки

Обзор

Lambowitz, A. M. and Belfort, M.(1993). Introns as mobile genetic elements. *Annu. Rev. Biochem.* 62, 587–622.

Lambowitz, A. M. and Zimmerly, S. (2004). Mobile group II introns. *Annu. Rev. Genet.* 38, 1–35.

Исследование

Dickson, L., Huang, H. R., Liu, L., Matsuura, M., Lambowitz, A. M., and Perlman, P. S. (2001). Retrotransposition of a yeast group II intron occurs by reverse splicing directly into ectopic DNA sites. *Proc. Natl. Acad. Sci. USA* 98, 13207–13212.

Zimmerly, S. et al. (1995). Group II intron mobility occurs by target DNA-primed reverse transcription. *Cell* 82, 545–554.

Zimmerly, S. et al. (1995). A group II intron is a catalytic component of a DNA endonuclease involved in intron mobility. *Cell* 83, 529–538.

27.7 Некоторым автокаталитическим интронам требуется матураза

Исследование

Bolduc et al. (2003). Structural and biochemical analyses of DNA and PHK binding by a bifunctional homing endonuclease and group I splicing factor. *Genes. Dev.* 17, 2875–2888.

Carignani, G. et al. (1983). An PHK maturase is encoded by the first intron of the mitochondrial gene for the subunit I of cytochrome oxidase in *S. cerevisiae*. *Cell* 35, 733–742.

Henke, R. M., Butow, R. A., and Perlman, P. S. (1995). Maturase and endonuclease functions depend on separate conserved domains of the bifunctional protein encoded by the group I intron aI4 alpha of yeast mitochondrial DNA. *EMBO J.* 14, 5094–5099.

Matsuura, M., Noah, J. W., and Lambowitz, A. M. (2001). Mechanism of maturase-promoted group II intron splicing. *EMBO J.* 20, 7259–7270.

27.9 Вироиды обладают каталитической активностью

Обзор

Doherty, E. A. and Doudna, J. A. (2000). Ribozyme structures and mechanisms. *Annu. Rev. Biochem.* 69, 597–615.

Symons, R. H. (1992). Small catalytic РНКs. *Annu. Rev. Biochem.* 61, 641–671.

Исследование

Forster, A. C. and Symons, R. H. (1987). Self-cleavage of virusoid РНК is performed by the proposed 55-nucleotide active site. *Cell* 50, 9–16.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983). The РНК moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* 35, 849–857.

Scott, W. G., Finch, J. T., and Klug, A. (1995). The crystal structure of an all-РНК hammerhead ribozyme: a proposed mechanism for РНК catalytic cleavage. *Cell* 81, 991–1002.

27.10 Редактирование РНК касается конкретных азотистых оснований

Исследование

Higuchi, M. et al. (1993). РНК editing of AMPA receptor subunit GluR-B: a base-paired intron-exon structure determines position and efficiency. *Cell* 75, 1361–1370.

Navaratnam, N. et al. (1995). Evolutionary origins of apoB mРНК editing: catalysis by a cytidine deaminase that has acquired a novel РНК-binding motif at its active site. *Cell* 81, 187–195.

Powell, L. M., Wallis, S. C., Pease, R. J., Edwards, Y. H., Knott, T. J., and Scott, J. (1987). A novel form of tissue-specific РНК processing produces apolipoprotein-B48 in intestine. *Cell* 50, 831–840.

Sommer, B. et al. (1991). РНК editing in brain controls a determinant of ion flow in glutamate-gated channels. *Cell* 67, 11–19.

27.11 Редактирование транскриптов направляется специальными малыми РНК

Исследование

Aphasizhev, R., Sbicego, S., Peris, M., Jang, S. H., Aphasizheva, I., Simpson, A. M., Rivlin, A., and Simpson, L. (2002). Trypanosome mitochondrial 3′ terminal uridylyl transferase (TUTase): the key enzyme in U-insertion/deletion РНК editing. *Cell* 108, 637–648.

Benne, R., Van den Burg J., Brakenhoff, J. P., Sloof, P., Van Boom, J. H., and Tromp, M. C. (1986). Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell* 46, 819–826.

Blum, B., Bakalara, N., and Simpson, L. (1990). A model for РНК editing in kinetoplastid mitochondria: "guide" РНК molecules transcribed from maxicircle DNA provide the edited information. *Cell* 60, 189–198.

Feagin, J. E., Abraham, J. M., and Stuart, K. (1988). Extensive editing of the cytochrome c oxidase III transcript in *Trypanosoma brucei*. *Cell* 53, 413–422.

Seiwert, S. D., Heidmann, S. and Stuart, K. (1996). Direct visualization of uridylate deletion *in vitro* suggests a mechanism for kinetoplastid editing. *Cell* 84, 831–841.

27.12 Сплайсинг белка — автокаталитическая реакция

Обзор

Paulus, H. (2000). Protein splicing and related forms of protein autoprocessing. *Annu. Rev. Biochem.* 69, 447–496.

Исследование

Derbyshire, V., Wood, D. W., Wu, W., Dansereau, J. T., Dalgaard, J. Z., and Belfort, M. (1997). Genetic definition of a protein-splicing domain: functional mini-inteins support structure predictions and a model for intein evolution. *Proc. Natl. Acad. Sci. USA* 94, 11466–11471.

Perler, F. B. et al. (1992). Intervening sequences in an Archaea DNA polymerase gene. *Proc. Natl. Acad. Sci. USA* 89, 5577–5581.

Xu, M. Q., Southworth, M. W., Mersha, F. B., Hornstra, L. J., and Perler, F. B. (1993). *In vitro* protein splicing of purified precursor and the identification of a branched intermediate. *Cell* 75, 1371–1377.

Хромосомы

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

28.1 Введение

28.2 Вирусный геном упакован в оболочку

- Структура капсида накладывает ограничения на длину ДНК, которая может быть встроена в вирусную частицу.
- Нуклеиновые кислоты внутри капсиды чрезвычайно конденсированны.
- У нитевидных РНК-вирусов конденсация РНК-генома происходит по мере сборки капсида вокруг него.
- У сферических ДНК-вирусов ДНК укладывается в заранее подготовленный белковый футляр.

28.3 Бактериальный геном образует нуклеоид

- Бактериальный нуклеоид, который по массе на 80% состоит из ДНК, можно разрыхлить, воздействуя на РНК или белок.
- Белки, ответственные за конденсацию ДНК, пока не идентифицированы.

28.4 Бактериальный геном сверхспирализован

- В нуклеоиде есть около 100 независимых доменов с отрицательной сверхспирализацией.
- Средняя плотность сверхспирализации составляет примерно 1 сверхвиток на 100 пн.

28.5 В ДНК эукариот есть петли и домены, прикрепленные к каркасу

- В интерфазном хроматине ДНК отрицательно сверхспирализована и распределена по независимым доменам в ~85 тпн.
- Метафазные хромосомы содержат белковый каркас, к которому прикреплены петли сверхспирализованной ДНК.

28.6 В интерфазе ДНК крепится к ядерному матриксу определенными нуклеотидными последовательностями

- ДНК прикреплена к матриксу определенными последовательностями, называемыми MAR или SAR.
- Последовательности MAR — это A-T-богатые регионы, но консенсуса не имеют.

28.7 Хроматин подразделяют на эухроматин и гетерохроматин

- Обособленные хромосомы можно наблюдать только при митозе.
- Во время интерфазы хроматин, по большей части, пребывает в форме эухроматина, который упакован гораздо менее плотно, чем митотические хромосомы.
- Участки конститутивного гетерохроматина остаются плотно упакованными в продолжение всей интерфазы

28.8 У хромосом проявляются полосатые паттерны окрашивания

- Определенные методики окрашивания придают хромосомам исчерченность в виде серий G-полос, или Гимза-дисков.
- В полосах содержание G-C ниже, чем в промежутках между ними.
- Гены сосредоточены в G-C-богатых светлых промежутках.

28.9 Хромосомы типа ламповых щеток имеют выпетливания

- У хромосом типа ламповых щеток сайты экспрессии генов выглядят как петли, выпущенные из оси хромосомы.

28.10 Политенные хромосомы выглядят полосатыми

- Окрашенные препараты политенных хромосом двукрылых насекомых несут серии полос, которые можно использовать как цитологическую карту.

28.11 Политенные хромосомы образуют «вздутия» в местах экспрессии генов

- На окрашенных препаратах политенных хромосом полосы, где шла активная экспрессия генов, выглядят раздутыми (имеют вид «пуфов»).

28.12 Эукариотические хромосомы служат клетке приспособлениями для сегрегации генетического материала

- Эукариотические хромосомы удерживаются в веретене деления благодаря микротрубочкам, прикрепленным к кинетохорам, формирующимся при центромерах.
- Во многих случаях центромеры содержат гетерохроматин, богатый последовательностями сателлитной ДНК.

28.13 Центромеры содержат повторяющиеся последовательности ДНК

- Центромеры в хромосомах высших эукариот содержат большие количества повторов ДНК.
- Функция этих повторов пока неизвестна.

28.14 Центромеры *S. cerevisiae* несут короткие последовательности ДНК

- Последовательности *CEN* у *S.cerevisiae* были идентифицированы как придающие плазмиде способность к правильной митотической сегрегации.
- Последовательности *CEN* состоят из коротких консервативных элементов *CDE-I* и *CDE-III*, которые фланкируют A-T-богатый элемент *CDE-II*.

28.15 *Центромера связывается с белковым комплексом*
- На *CDE-II* формируется специализированный белковый комплекс, альтернативный обычной структуре хроматина.
- С *CDE-III* связывается белковый комплекс CBF3.
- Белки, соединяющие эти два комплекса, могут обеспечивать их связь с микротрубочками.

28.16 Теломеры содержат простые повторяющиеся последовательности ДНК
- Теломера нужна для стабильности конца хромосомы.
- Теломерная ДНК состоит из многократного простого повтора, одна из цепей которого имеет вид $C_{>1}(A/T)_{1-4}$.

28.17 Теломеры опечатывают концы хромосом
- Белок TRF2 катализирует реакцию, в ходе которой 3′-концевой повтор G-хвоста замещает гомологичный повтор, вытеснив его из дуплекса в вышележащей области теломерной ДНК.

28.18 *Теломеры синтезируются* рибонуклеопротеиновым ферментом
- Теломераза использует 3′-конец G-хвоста теломеры как затравку для синтеза тандемных повторов TTGGGG.
- РНК-компонент теломеразы обладает последовательностью, идентичной C+A-богатым повторам теломерной ДНК.
- Каталитическая субъединица теломеразы – это обратная транскриптаза, использующая РНК-компонент как матрицу для синтеза G+T-богатой последовательности теломерной ДНК.

28.19 Теломеры принципиально важны для сохранности хромосом и выживания клеточных линий

28.20 Резюме

28.1 Введение

В том, как организован генетический материал любой клетки, виден общий принцип. Этот материал существует в виде компактной массы в пределах ограниченного объема. При этом разнообразные активности, такие как репликация и транскрипция, должны иметь место внутри отведенного ему пространства. Организация генетического материала должна быть приспособлена к его переходам из неактивного состояния в активное состояние, и обратно.

Конденсированное состояние нуклеиновой кислоты — это результат ее взаимодействия с белками, которые, с химической точки зрения, обладают основными свойствами. Положительные заряды этих белков нейтрализуют отрицательные заряды нуклеиновой кислоты. Структура нуклеопротеинового комплекса определяется взаимодействием белков с ДНК (или РНК).

Размещение ДНК в частицах вирусов и фагов, клетках бактерий и ядрах эукариот наталкивается на общую проблему. Во всех случаях протяженность ДНК, во всю ее молекулярную длину, весьма и весьма превосходит размеры компартмента, в котором она содержится. ДНК (или, у некоторых вирусов, РНК) должна быть чрезвычайно плотно спрессована, чтобы поместиться в отведенное ей пространство. *Таким образом, вразрез с привычным представлением о ДНК, как протяженной двойной спирали, структурная деформация ДНК (сгибание, складывание) с целью придания ей более компактной формы — это скорее правило, чем исключение.*

Насколько велико несоответствие между длиной нуклеиновой кислоты и размером занимаемого ею компартмента, можно судить из таблицы на рис. 28.1. У бактериофагов и вирусов эукариот геном, состоит ли он из одно- или двухцепочечной, ДНК или РНК, эффективно заполняет свой капсид (который может быть палочковидным или сферическим).

Что же касается бактерий или клеточных компартментов эукариот, степень сжатия ДНК в них трудно вычислить точно, поскольку ДНК содержится в компактной зоне, занимающей лишь часть компартмента. Генетичес-

В условиях *in vivo* любая геномная ДНК очень компактна

Компартмент	Его форма	Его размеры	Нуклеиновая кислота	Ее длина
ВТМ	филамент	0,008 × 0,3 мкм	Одна одноцепочечная молекула РНК	2 мкм = 6,4 тпн
Фаг fd	филамент	0,006 × 0,85 мкм	Одна одноцепочечная молекула ДНК	2 мкм = 6,0 тпн
Аденовирус	икосаэдр	0,07 мкм в диаметре	Одна двуцепочечная молекула ДНК	11 мкм = 35,0 тпн
Фаг T4	икосаэдр	0,065 × 0,10 мкм	Одна двуцепочечная молекула ДНК	55 мкм = 170,0 тпн
E. coli	цилиндр	1,7 × 0,65 мкм	Одна двуцепочечная молекула ДНК	1,3 мм = $4{,}2 \times 10^3$ тпн
Митохондрия (человек)	вытянутый шар	3,0 × 0,5 мкм	~10 одинаковых двуцепочечных молекул ДНК	50 мкм = 16,0 тпн
Клеточное ядро (человек)	шар	6 мкм в диаметре	46 хромосом из двуцепочечной ДНК	1,8 м = 6×10^6 тпн

Длина молекулы нуклеиновой кислоты во много раз больше, чем размеры компартмента, в который она заключена

кий материал можно увидеть как **нуклеоид** в бактериях и как массу **хроматина** в ядрах эукариот во время интерфазы (в период между клеточными делениями).

Плотность ДНК в этих компартментах очень высока. В бактерии она составляет около 10 мг/мл, в эукариотическом ядре — около 100 мг/мл, а в головке фага T4 — более 500 мг/мл. Такая концентрация ДНК в растворе была бы эквивалентна гелю огромной вязкости. Физиологический смысл этого обстоятельства (например, как это влияет на способность белков находить свои сайты связывания на ДНК) пока не исследован.

Упаковка хроматина отличается гибкостью; она меняется по ходу эукариотического клеточного цикла. Во время деления (митоза или мейоза) генетический материал становится упакованным еще более плотно, так что становятся узнаваемы отдельные **хромосомы**.

Общее сжатие ДНК может быть выражено как **коэффициент упаковки**, представляющий собой длину молекулы ДНК, деленную на длину структурной единицы, которая ее содержит. Например, самая маленькая из хромосом человека содержит примерно $4,6 \times 10^7$ пн ДНК (что примерно в 10 раз превышает размер генома бактерии *E. coli*). Это соответствует 14000 мкм (1,4 см) двойной спирали ДНК. В самый «конденсированный» момент митоза эта самая маленькая хромосома имеет длину около 2 мкм. Таким образом, коэффициент упаковки ДНК в хромосоме человека может доходить до 7000.

Для более аморфных структур (бактериального нуклеоида и эукариотического хроматина) коэффициенты упаковки ДНК не могут быть установлены с такой точностью. Принято считать, что митотические хромосомы упакованы в 5–10 раз плотнее, чем интерфазный хроматин, так что характерное значение коэффициента упаковки попадает в интервал между 1000 и 2000.

Наконец, главный вопрос, который пока остается без ответа, касается *специфичности* упаковки. Подчинена ли укладка ДНК *конкретному* паттерну, или индивидуальна для каждой отдельно взятой копии генома? Как меняется паттерн упаковки, когда какой-нибудь сегмент ДНК участвует в транскрипции или репликации?

28.2 Вирусный геном упакован в оболочку

Ключевые понятия:

- Структура капсида накладывает ограничения на длину ДНК, которая может быть встроена в вирусную частицу.
- Нуклеиновые кислоты внутри капсида чрезвычайно конденсированны.
- У нитевидных РНК-вирусов конденсация РНК-генома происходит по мере сборки капсида вокруг него.
- У сферических ДНК-вирусов ДНК укладывается в заранее подготовленный белковый футляр.

С точки зрения упаковки *отдельно взятой* молекулы нуклеиновой кислоты, между клеточным геномом и вирусом есть существенная разница. Размер клеточного генома не постоянен; число и местоположение отдельно взятых последовательностей может быть изменено дупликациями, делециями и перестройками. В таком случае необходим инвариантный метод упаковки ДНК, нечувствительный к общему содержанию или распределению последовательностей. Нужды вируса, напротив, определяются двумя ограничениями. Размер нуклеиновой кислоты, предназначенной к упаковке, *предопределен* размером вирусного генома; он обязан поместиться внутри оболочки, собранной из белков, кодируемых генами вируса.

Устройство вирусной частицы может показаться простым, но это впечатление обманчиво. Геном — нуклеиновая кислота — находится внутри **капсида**, представляющего собой симметричную или квазисимметричную структуру, собранную из белковых субъединиц одной или двух-трех разновидностей. К капсиду прикреплены (или вмонтированы в него) другие структуры; они собраны из особых белков и необходимы для инфицирования будущего хозяина.

Вирусная частица обладает очень плотной структурой. Внутренний объем капсида редко бывает больше объема нуклеиновой кислоты, которую он должен хранить. Обычно нуклеиновая кислота занимает более половины внутреннего объема капсида, а нередко — и весь объем.

То условие, что капсид обязан быть собран из белков, кодируемых вирусом, в пределе означает, что оболочка будет целиком сконструирована из однотипных субъединиц. Согласно правилам сборки идентичных субъединиц в замкнутые структуры, геометрическая форма капсида при этом сведется к одному из двух вариантов. В первом случае белковые субъединицы последовательно укладываются в спиралевидную решетку, образуя капсид *нитевидной*, или палочковидной, формы. Во втором случае, одинаковые субъединицы могут собраться в псевдосферическую оболочку — тип структуры, удовлетворяющий определению многогранника с **икосаэдрической симметрией**. Капсиды некоторых вирусов собираются из более чем одного типа белковых субъединиц, но, хотя это и вносит некоторые уточнения в типы структур, которые могут быть сформированы, тем не менее, все вирусные капсиды образованы либо квазикристаллическими филаментами, либо икосаэдрами.

Существует два возможных варианта посроения капсида:

- Капсид может быть собран вокруг нуклеиновой кислоты, конденсируя при этом ДНК или РНК за счет взаимодействия белка и нуклеиновой кислоты в процессе сборки.
- Капсид может быть сконструирован как полая оболочка, которую должна заполнить нуклеиновая кислота, конденсируясь по мере своего туда вхождения.

У одноцепочечных РНК-вирусов капсид собирается вокруг генома. Принцип сборки заключается в том, что *положение РНК внутри капсида напрямую определя-*



РИС. 28.2 В частице (вирионе) ВТМ геномная РНК идет по спирали, поддерживаемая стопкой дисков из одинаковых белковых субъединиц

ется ее связыванием с белками оболочки. Самый известный тому пример — это вирус табачной мозаики (ВТМ; tobacco mosaic virus, TMV). Сборка начинается на месте двухцепочечной шпильки, заложенной внутри последовательности РНК. Из этого центра зарождения (**нуклеации**, nucleation center) сборка распространяется в обоих направлениях вдоль РНК, пока не дойдет до ее концов. Сборочной единицей капсида является двуслойный диск, в котором каждый слой содержит 17 одинаковых белковых субъединиц. Диск — это цилиндрическая структура, которая при взаимодействии с ДНК приобретает форму спирали. Самый первый диск меняет свою конформацию, когда в его отверстие входит шпилька РНК. Каждый последующий шаг сборки добавляет новый диск, который затягивает в свое центральное отверстие новый отрезок РНК. В результате РНК оказывается уложенной в виде спирали на внутренней стороне белковой оболочки, как показано на РИС. 28.2.

Сборка сферических капсидов ДНК-вирусов происходит по-другому; подробнее всего это описано для фагов лямбда и Т4. Из небольшого набора белков собирается пустой капсид. *Затем в него входит двухцепочечный геном вируса,* что сопровождается изменением структуры капсида.

На РИС. 28.3 показана сборка вириона лямбда. Капсид закладывается как белковая сердцевина, которая превращается в полую оболочку более четкой формы. В этот момент начинается упаковка ДНК, и оболочка увеличивается в размерах, сохраняя при этом свою характерную форму. Наконец, наполненная головка запечатывается добавлением хвоста.

Двухцепочечная ДНК, если ее рассматривать на коротком участке, подобна довольно жесткой палке. И все же она должна быть как-то сжата в компактную структуру, чтобы поместиться в капсид. Какого рода деформации обеспечивают укладку ДНК в капсид: равномерное скручивание или резкое изгибание?

В укладке ДНК в капсид фага задействованы два типа реакций: транслокация и конденсация. Обе они энергетически невыгодны.

Транслокация — это активный АТФ-зависимый процесс, увлекающий ДНК внутрь капсида. Механизм используется многими вирусами, которые реплицируются по типу «катящегося кольца» с образованием длинных «хвостов», содержащих мультимеры вирусного генома. Лучше всего это описано у фага лямбда. Его геном упаковывается в пустой капсид ферментом **терминазой**. Процесс показан на РИС. 28.4.

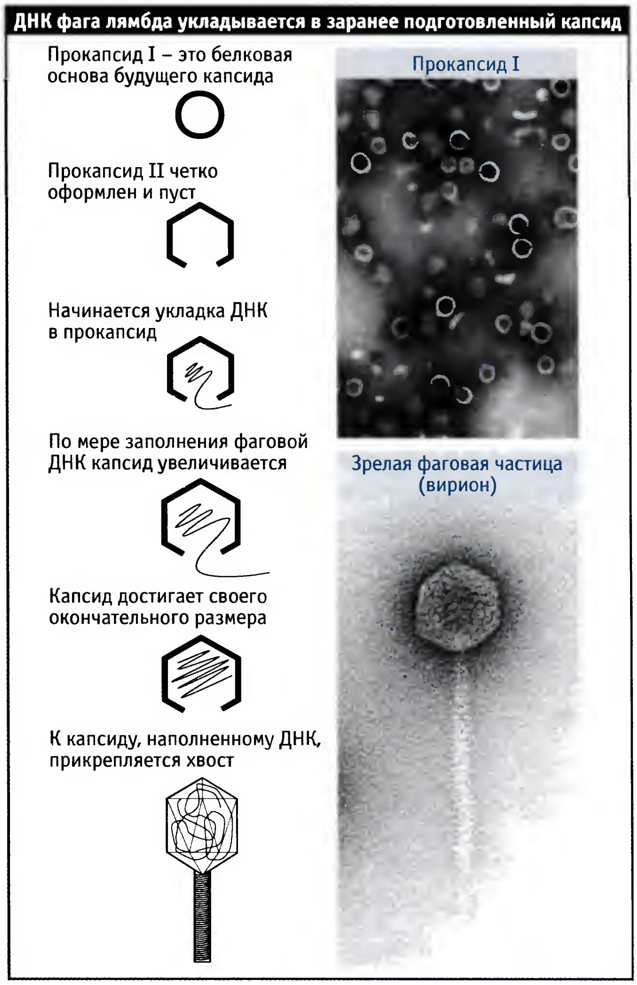


РИС. 28.3 Созревание фага лямбда проходит несколько стадий. Пустой капсид (прокапсид) заполняется фаговой ДНК; при этом он меняет свою форму и увеличивается в размере. На фото запечатлены фаговые частицы в начале и в конце пути созревания. Верхнее фото воспроизведено по Cue, D. and Feiss, M. 1993. Proc. Natl. Acad. Sci. USA. 90: 9290–9294. Авторские права 1993 National Academy of Science, U.S.A. Фотография любезно предоставлена Michael G. Feiss, University of Iowa. Нижняя фотография любезно предоставлена Robert Duda, University of Pittsburgh

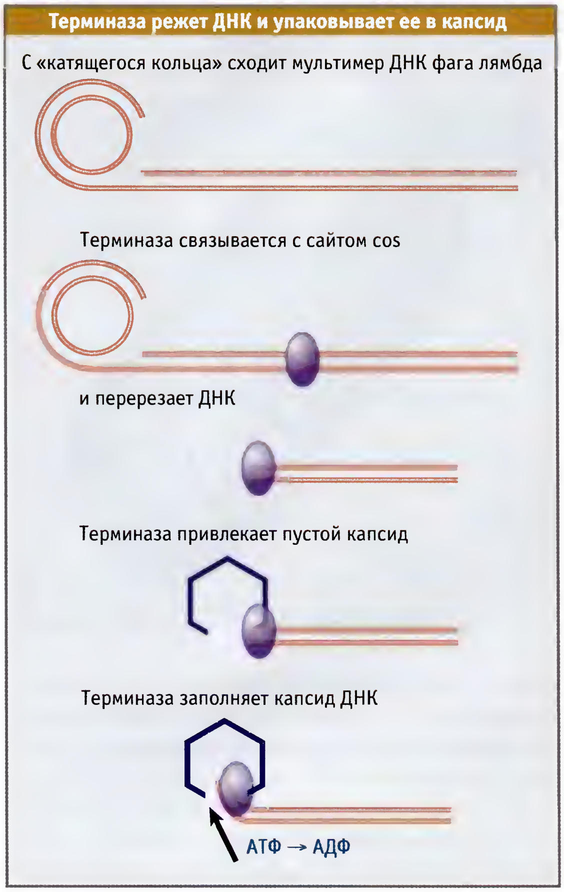


РИС. 28.4 Белок терминаза связывается с конкретными сайтами в мультимере генома лямбда – продукте репликации по типу «катящегося кольца». Терминаза отрезает единичный фрагмент ДНК (полноразмерную копию генома лямбда) и связывается с пустым капсидом. После этого терминаза, используя энергию гидролиза АТФ, заполняет капсид ДНК

О терминазе вначале узнали в связи с ее ролью в образовании концов линейной ДНК фага: она разрезает мультимер по сайтам *cos*. (Название *cos* отражает тот факт, что действие фермента порождает липкие (cohesive) концы, имеющие комплементарные одноцепочечные выступы.) Геном фага кодирует две субъединицы, вместе составляющие терминазу. Одна субъединица связывается с сайтом *cos*; после чего к ней присоединяется другая субъединица, разрезающая ДНК. Терминаза собирается в гетероолигомер в составе комплекса, в котором, кроме нее, состоит хозяйский фактор интеграции, димер, кодируемый геномом бактерии. Затем терминаза связывается с пустым капсидом и использует гидролиз АТФ, чтобы получить энергию, необходимую ей для транслокации (смещения) вдоль ДНК. Транслокация терминазы заправляет ДНК в капсид.

В других случаях транслокация может обеспечиваться структурным компонентом бактериофага. У ϕ29 *Bacillus subtilis* мотор, загоняющий ДНК в капсид — это структура, соединяющая головку и хвост. Она работает как поворотный двигатель с эффектом линейного перемещения ДНК внутрь капсида. Когда фаг инфицирует бактерию, тот же мотор используется для выброса ДНК из капсида.

О механизме, заставляющем ДНК конденсироваться внутри пустого капсида, известно мало. Во всяком случае, кроме ДНК капсид содержит «внутренние белки». Возможно, они обеспечивают своего рода каркас, на котором конденсируется ДНК (что было бы аналогично использованию белков в оболочке РНК-вирусов растений).

Насколько специфичен процесс упаковки у вирусов? Его зависимость от конкретных последовательностей исключена, поскольку мутации (делеции, вставки или замены) в геноме вируса процессу сборки не мешают. Взаимодействие между капсидом и ДНК исследовали напрямую — определением участков ДНК, которые могут быть химически связаны с белками капсида. Неожиданно оказалось, что все участки ДНК взаимодействуют с капсидом более или менее равновероятно. Это может означать, что, когда ДНК входит в капсид, она конденсируется согласно некому общему правилу, и паттерн конденсации не обусловлен никакими особенными последовательностями.

Разнообразные механизмы сборки вирусной частицы приводят к одному и тому же финалу: одиночной молекуле ДНК или РНК, упакованной в капсид. Впрочем, некоторые вирусы имеют геномы, состоящие из нескольких нуклеиновых кислот. Реовирус содержит десять двухцепочечных сегментов РНК. Чтобы обеспечить каждой частице гарантированную выборку из одной копии каждого образца (то есть полный объем генетической информации), используются специфические сортировочные последовательности сегментов РНК. В более простом случае, у фага ϕ6, в капсид упаковывается три различных сегмента двухцепочечной РНК. Эти сегменты должны войти туда в особом «регламентированном» порядке: вход сегмента срабатывает на изменение конформации капсида, которое создает сайты связывания со следующим сегментом.

Некоторые вирусы растений состоят из нескольких частиц: их геномы состоят из сегментов, каждый из которых упакован в *отдельный* капсид. Например, у вируса мозаики люцерны (ВМЛ) четыре разных одноцепочечных РНК независимо упаковываются каждая в отдельную оболочку из одних и тех же белковых субъединиц. Успех инфекции зависит от попадания в клетку хотя бы одной частицы каждого типа.

Четыре типа частиц ВМЛ имеют разные размеры. Во всех четырех случаях один и тот же тип белковых субъединиц капсида складывается в отдельную форму, характерную для конкретного типа РНК, которую он упаковывает. Это обстоятельство представляет собой отступление от свойственного вирусам норматива упаковки нуклеотидной последовательности заданной длины в капсид фиксированной формы.

Путь сборки вирусов с капсидом единственно аутентичной формы может быть нарушен мутациями,

приводящими к образованию аберрантных частиц-**монстров** с более длинной головкой, чем в норме. Это означает, что, несмотря на присущую белкам капсида способность к сборке в структуру конкретного типа, точные размеры и форма этой структуры могут варьировать. Некоторые из таких мутаций находятся в генах, кодирующих **факторы сборки**, необходимые для формирования головки частицы. При этом сами факторы сборки в состав капсида не входят, но как вспомогательные белки ограничивают сборочные опции для субъединиц белка капсида и направляют процесс сборки на путь, единственно приемлемый для нормальной частицы. Белки с похожей функцией задействованы при упаковке хроматина в эукариотической клетке (см. гл. 29 «Нуклеосомы»).



РИС. 28.5 На тонком срезе нуклеоид выглядит как плотная масса в центре клетки. Фотография любезно предоставлена the Molecular and Cell Biology Instructional Laboratory Program, University of California, Berkeley

28.3 Бактериальный геном образует нуклеоид

Ключевые понятия:

- Бактериальный нуклеоид, который по массе на 80% состоит из ДНК, можно разрыхлить, воздействуя на РНК или белок.
- Белки, ответственные за конденсацию ДНК, пока не идентифицированы.

Хотя у бактерий и нет структур с четкой морфологией эукариотических хромосом, их геномы, несомненно, структурно организованны и имеют вид плотных телец. На электронных микрофотографиях генетический материал бактерий можно увидеть как достаточно плотный комок (или скопление комков), занимающий около трети объема бактериальной клетки. На РИС. 28.5 показан тонкий срез, где хорошо различим нуклеоид.

Если клетки *E. coli* подвергнуть лизису, то фибриллы ДНК высвобождаются в виде петель, прикрепленных к поврежденной оболочке клетки. Как можно видеть на РИС. 28.6, ДНК в составе этих петель не расправлена в свободный дуплекс, но сохраняет упакованное состояние вследствие своей ассоциации с белками.

У *E. coli* было обнаружено несколько белков, отдаленно напоминающих хромосомные белки эукариот. Попробуем перечислить критерии, которыми следует руководствоваться, решая, может ли данный ДНК-связывающий белок играть структурную роль в нуклеоиде. Прежде всего, он должен присутствовать в количествах, которые были бы достаточны для его связывания на всем протяжении генома. Кроме того, мутации в его гене должны приводить к какому-нибудь структурному или функциональному нарушению, которое угрожало бы выживанию генома (например, сегрегации по дочерним клеткам), и этому генетическому критерию пока что не удовлетворяет ни один из белков-кандидатов.

Белок HU — это димер, конденсирующий ДНК (возможно, путем упаковки в виде бус). HU похож на хозяйский фактор интеграции, играющий структурную роль в сайт-специфичной рекомбинации. Нуль-мутации в любом из генов, кодирующих субъединицы HU (*hupA* и *hupB*) имеют слабый эффект, однако потеря клеткой



РИС. 28.6 Из лизированной клетки *E. coli* «выскочил» нуклеоид, похожий на петли длинного волокна. Фотография © G. Murti/Photo Researchers, Inc.

функций обеих субъединиц сопряжена с фенотипом чувствительности к низким температурам и некоторым уменьшением сверхспирализованности ДНК. Эти данные говорят в пользу общего значения HU для конденсации нуклеоида.

Белок H1 (известный также как H-NS) связывает ДНК (предпочтительно — изогнутые последовательности). Мутации в его гене имеют различные фенотипические проявления (*osmZ*, *bglY*, *pilG*), словно принадлежа регуляторам различных бактериальных клеточных систем. В таком многообразии фенотипов, возможно, сказывается влияние H1 на локальную топологию ДНК, и через нее — на экспрессию различных генов в манере, зависящей от конкретного промотора.

Почему же фенотипические эффекты делеций в генах белков HU и H1 сравнительно мягки? Ведь от отсутствия белка, структурирующего нуклеоид, можно было бы ожидать серьезной угрозы для выживания бактериальной клетки? Одно из объяснений состоит в том, что в нормальной клетке эти белки дублируют друг друга, и любой из них может восполнить потерю другого. Поэтому, чтобы всерьез нарушить структуру нуклеоида, необходимы делеции во *всех* соответствующих генах. Возможно, идентифицировать белки, определяющие главные структурные черты, равно как и целостность нуклеоида, еще только предстоит.

Нуклеоид может быть изолирован в форме быстро седиментирующего комплекса, по массе на ~80% состоящего из ДНК. (В аналогичных комплексах эукариот массовая доля ДНК равна ~50%; см. разд. 28.4 «Бактериальный геном сверхспирализован».) Нуклеоид как структуру, можно распустить обработкой веществами, действующими на РНК или на белок. Возможная роль белков в стабилизации нуклеоида как структуры очевидна. Роль РНК пока не выяснена.

28.4 Бактериальный геном сверхспирализован

Ключевые понятия:

- В нуклеоиде есть около 100 независимых доменов с отрицательной сверхспирализацией.
- Средняя плотность сверхспирализации составляет примерно 1 сверхвиток на 100 пн.

ДНК бактериального нуклеоида, помещенная в условия *in vitro*, ведет себя как замкнутая двухцепочечная структура. Такой вывод можно сделать на основании ее взаимодействия с бромистым этидием. Небольшая молекула бромистого этидия интеркалирует (вклинивается) между парами оснований с образованием *положительных* сверхвитков в замкнутых кольцевых молекулах ДНК, то есть в молекулах, у которых обе цепи дуплекса обладают ковалентной целостностью. (Во «вскрытых» кольцевых молекулах с надрезом по одной из цепей — как и в линейных молекулах — ДНК может ответить на интеркаляцию бромистого этидия свободным вращением, ослабляющим возникшее напряжение.)

В ДНК, замкнутой от природы и сверхспирализованной *отрицательно*, интеркаляция бромистого этидия вначале устраняет отрицательные сверхвитки, а затем приводит к появлению положительных сверхвитков. Количество бромистого этидия, необходимое для достижения нулевой сверхспирализации, служит мерой исходной плотности отрицательных сверхвитков.

При лабораторном выделении компактного нуклеоида из бактериальных клеток в его ДНК возникают нечаянные одноцепочечные разрывы; такие же разрывы можно получить при помощи щадящей обработки нуклеоида ДНКазой. Разрывы эти, однако, не отменяют появления положительных сверхвитков ДНК нуклеоида в ответ на обработку бромистым этидием. Это означает, что геномная ДНК бактерий, должно быть, структурирована в виде многочисленных независимых хромосомных **доменов**, так что *сверхспирализованность ДНК внутри каждого домена не нарушается событиями в других доменах*.

Такая локальная автономность позволяет приписать бактериальной хромосоме структуру, схематично изображенную на РИС. 28.7. Каждый домен состоит из петли ДНК, основание которой каким-то (неизвестным) способом закреплено, что не позволяет вращательным событиям распространяться из одного домена в другой.

В ранних работах размер домена оценивали как ~40 тпн ДНК, но сравнительно недавний анализ показал, что размеры доменов, возможно, меньше и равняются



РИС. 28.7 Геномная ДНК бактерии состоит из большого числа петель двухцепочечной ДНК (упакованной как волокно). Каждая петля фиксирована у основания и образует независимый структурный домен

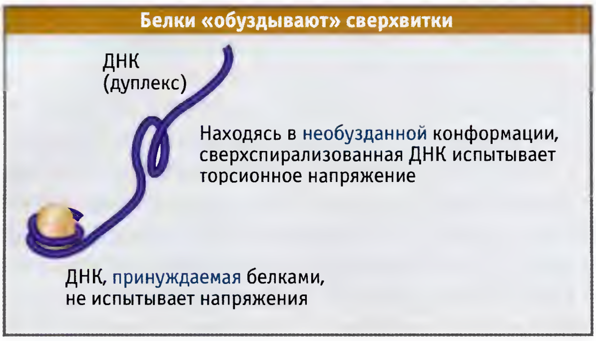


РИС. 28.8 Необузданный сверхвиток ДНК создает напряжение, которое распространяется вдоль молекулы. Если ДНК образует сверхвиток «по принуждению» белка, напряжения не возникает

~10 тпн. Это соответствует ~400 доменам на геном *E. coli*. Границы доменов, по всей видимости, распределены на хромосоме случайным образом, вне привязки к специализированным сайтам.

Существование обособленных доменов может способствовать подержанию разной степени сверхспирализации ДНК разных участков генома. Это соображение заставляет вспомнить, что промоторы бактерий дифференциальны по своей чувствительности к сверхспирализации (см. разд, 11.15 «Степень сверхспирализации ДНК влияет на процесс транскрипции»).

Как показано на РИС. 28.8, сверхспирализация геномной ДНК может, в принципе, принять одну из двух форм:

- Если сверхспирализованная ДНК свободна от белков, ее пространственная конфигурация *необуздана*, а наличие отрицательных сверхвитков порождает торсионное напряжение, свободно передаваемое вдоль ДНК. Напряжение может быть ослаблено расплетанием двойной спирали, как описано в разд. 19.12 «Сверхспирализация затрагивает структуру ДНК». ДНК пребывает в динамическом равновесии между состояниями напряжения и расплетания.
- Если ДНК связана с белками, удерживающими ее в *принудительной* пространственной конформации, сверхвитки представлены траекторией, которой ДНК следует, будучи жестко ассоциирована с белками. Энергия взаимодействия между белками и сверхспирализованной ДНК стабилизирует нуклеиновую кислоту, и никакого напряжения вдоль молекулы не передается.

Насколько необузданны *in vivo* сверхвитки геномной ДНК *E. coli*? Иными словами, какие торсионные напряжения, свойственные свободной ДНК, испытывает ее двойная спираль? Попытки измерить сверхспирализованность *in vitro* столкнулись с трудностью: невозможно выделить нуклеоид так, чтобы белки, которые влияют на конфигурацию ДНК, при этом не терялись. С помощью разных подходов было установлено, что ДНК испытывает торсионный стресс *in vivo*.

Один из подходов состоит в измерении эффекта от внесения в ДНК одноцепочечных разрывов. Необузданные сверхвитки при этом сходят на нет, а принудительные остаются без изменений. Внесение одноцепочечных разрывов сводит на нет ~50% сверхвитков. Это говорит о том, что примерно половина общего количества сверхспирализации передается вдоль ДНК как напряжение, а вторая половина абсорбируется через связывание белков.

В другом подходе используется сшивающий агент псорален. Это вещество проявляет большее сродство к ДНК, когда та находится в торсионном напряжении. Реакция псоралена с ДНК *E. coli* в условиях *in vivo* соответствует средней плотности в один отрицательный сверхвиток на 200 пн ($\sigma = -0{,}05$).

Также можно проверить клетки на способность к образованию альтернативных структур ДНК, например крестовидных пар шпилек в местах палиндромов. Исходя из изменения порядка зацепления, необходимого для осуществления таких реакций, можно вычислить исходную плотность сверхвитков. Этот метод выдает среднюю плотность сверхвитков $\sigma = -0{,}025$, или один отрицательный сверхвиток на 100 пн.

Таким образом, сверхвитки *создают* торсионное напряжение *in vivo*. Его средний уровень может варьировать, и точно измерить разброс плотностей крайне трудно. И все же ясно, что напряжения такого уровня хватает для оказания ощутимых воздействий на структуру ДНК — например, содействия плавлению в областях ориджинов или промоторов.

Многие важные аспекты устройства компактного нуклеоида остаются нераскрытыми. По какому принципу организуются домены? Лежат ли одни и те же последовательности — относительно разбиения на домены — в одних и тех же местоположениях? Могут ли границы доменов сдвигаться? Каким образом поддерживается целостность доменов? Биохимический анализ, сам по себе, не может дать исчерпывающих ответов на эти вопросы; если окажется возможной разработка подходящих селективных методик, то свойства структурных мутантов послужат наводкой в молекулярном анализе строения нуклеоида.

28.5 В ДНК эукариот есть петли и домены, прикрепленные к каркасу

Ключевые понятия:

- В интерфазном хроматине ДНК отрицательно сверхспирализована и распределена по независимым доменам в ~85 тпн.
- Метафазные хромосомы содержат белковый каркас, к которому прикреплены петли сверхспирализованной ДНК.

В контрасте с высокоорганизованной и четко воспроизводимой ультраструктурой митотических хромосом,



РИС. 28.9 Препараты хромосом, обедненные по гистонам, состоят из белкового каркаса, на котором закреплены петли ДНК. Воспроизводится по Cell, vol. 12, Paulson, J. R., and Laemmli, U. K., The structure of histone . . . , pp. 817–828. Авторские права 1977, с разрешения Elsevier. Фотография любезно предоставлена Ulrich K. Laemmli, University of Geneva, Switzerland

интерфазный хроматин представляет собой спутанную массу, занимающую значительную часть объема клеточного ядра. Чем контролируется распределение интерфазного хроматина внутри ядра?

Косвенным образом, природа этого распределения приоткрывается при лабораторном выделении генома в виде единого компактного тела. Пользуясь той же методикой, что была разработана для выделения бактериального нуклеоида (см. разд. 28.4 «Бактериальный геном сверхспирализован»), ядра можно подвергнуть лизису в градиенте сахарозы. Эта процедура высвобождает геном в такой форме, что его можно затем отделить центрифугированием. Выделенный геном *Drosophila melanogaster* может быть визуализирован как компактно сложенная нить (волокно, фибрилла) диаметром 10 нм, состоящая из ДНК, связанной с белками.

Сверхспирализованность эукариотической геномной ДНК, измеренная путем обработки бромистым этидием, равняется примерно одному отрицательному сверхвитку на 200 пн. Эти сверхвитки могут быть убраны ДНКазой (то есть внесением одноцепочечных разрывов), причем ДНК сохранит свою форму фибриллы диаметром 10 нм. Это говорит о том, что сверхспирализация — это результат принудительного размещения волокна в пространстве.

Для полной релаксации сверхвитков требуется внесение одного ника на 85 тпн, которые, таким образом, соответствуют средней длине замкнутой ДНК. Замкнутые участки могут представлять собой петли или домены, подобные тем, что были обнаружены в геномной ДНК бактерий. Петли можно увидеть, если экстрагировать из митотических хромосом бóльшую часть белков. Комплекс, который при этом останется, будет состоять из ДНК, ассоциированной с 8% исходного содержания белков. На РИС. 28.9 показано, как хромосомы, обедненные белками, принимают форму центрального **каркаса** (scaffold), окруженного ореолом ДНК.

Каркас метафазной хромосомы состоит из плотно переплетенных волокон. Нити ДНК выбиваются из каркаса в виде петель со средней длиной от 10 до 30 мкм (то есть от 30 до 90 тпн). ДНК можно расщепить, не нарушив целостности каркаса, состоящего из набора специальных белков. По всей видимости, петли ДНК заякорены в центральном каркасе из белков.

Внешний вид каркаса напоминает митотическую пару сестринских хроматид. Обычно сестринские каркасы тесно соприкасаются (хотя иногда выглядят разделенными) и соединены всего лишь несколькими волокнами. Может ли такая структура быть ответственна за поддержание формы митотических хромосом? Может ли каркас возникать как средоточие компонентов интерфазного хроматина, которыми скреплены основания его петель?

28.6 В интерфазе ДНК крепится к ядерному матриксу определенными нуклеотидными последовательностями

Ключевые понятия:

- ДНК прикреплена к матриксу определенными последовательностями, называемыми MAR или SAR.
- Последовательности MAR — это A-T-богатые регионы, но консенсуса не имеют.

Возможно ли, что ДНК прикрепляется к каркасам хромосом какими-то особыми последовательностями? На срезах интерфазных ядер хроматин часто выглядит прикрепленным к матриксу, и не раз высказывалось предположение, что это прикрепление необходимо для транскрипции или репликации. Сайты ДНК, прикрепленные к белковым структурам в интерфазных ядрах, называют **MAR** (matrix attachment regions; участки прикрепления к матриксу), иногда их также называют *SAR* (scaffold attachment regions; участки прикрепления к каркасу). Что представляет собой структура, с которой они соединены, не ясно. Частичное освобождение ядра от белков приводит к выпетливанию ДНК из остаточной белковой структуры. Попытки соотнести белки, обнаруживаемые в таких препаратах, со структурными элементами интактных клеток успехом не увенчались.

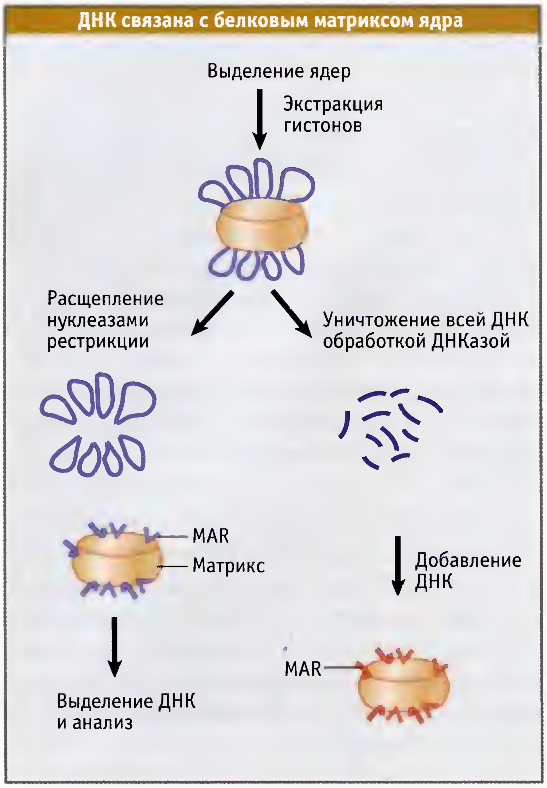


РИС. 28.10 Участки ДНК, ассоциированные с белками ядерного матрикса, можно выявить, «срезав» с матрикса свободную ДНК. Обратный экспериментальный подход заключается в идентификации таких участков по их высокому сродству к очищенному матриксу *in vitro*

Какие именно участки ДНК ассоциированы с матриксом? Экспериментальные подходы к ответу на этот вопрос иллюстрирует РИС. 28.10. Оба подхода начинаются с выделения суммарного ядерного экстракта, содержащего хроматин и ядерные белки. К суммарному ядерному экстракту применяют различные варианты обработки, нацеленные на идентификацию дезоксирибонуклеотидных последовательностей, которые либо оказались прикрепленными к матриксу на момент его выделения (*in vivo*), либо способны прикрепиться к нему (*in vitro*).

Чтобы вычленить реально существующие MAR, хроматиновые петли деконденсируют (разрыхляют) экстракцией белков. После удаления ДНК-содержащих петель при помощи обработки эндонуклеазами рестрикции *in vivo* остаются (презумптивные) MAR, прикрепленные к матриксу.

Другой подход, комплементарный по смыслу, заключается в удалении из матрикса *всей* ДНК (посредством обработки ДНКазой), после чего любые отдельно взятые фрагменты ДНК можно проверить на способность к связыванию с матриксом *in vitro*.

По идее, последовательности, ассоциированные с матриксом *in vivo* и *in vitro* — одни и те же. Как только обнаруживается потенциальный MAR, размер его минимального участка, достаточного для ассоциации с матриксом *in vitro*, может быть определен при помощи делеций. Наконец, все тот же эксперимент дает возможность идентифицировать белки, находящиеся в контакте с MAR.

Последовательности MAR, вопреки ожиданиям, не консервативны. Обычно они A-T-богаты (~70%), но консенсуса у них нет. Нередко в окрестностях MAR можно обнаружить разные интересные последовательности: например, сайты *цис*-регуляции транскрипции. Обычно MAR содержат сайт узнавания для топоизомеразы II, и вполне возможно, что, помимо прикрепления ДНК к матриксу, MAR служат и другим функциям, поскольку содержат сайты, по которым идет внесение изменений в топологию ДНК.

Какова взаимосвязь между каркасами хромосом делящихся клеток и матриксом интерфазных ядер? Прикреплены ли к ним одни и те же последовательности ДНК? В ряде случаев с каркасов метафазных хромосом удавалось собрать последовательности, идентичные находившимся *in vivo* в контакте с интерфазным матриксом. Фрагменты ДНК, содержащие последовательности MAR, могут связаться с каркасом метафазной хромосомы, так что, похоже, единый тип сайтов прикрепления существует. В интерфазе такие сайты крепятся к ядерному матриксу, а при митозе — к каркасам хромосом.

Ядерный матрикс и каркасы хромосом состоят из разных белков, хотя некоторые компоненты у них общие. Немаловажный компонент каркаса, топоизомераза II, входит и в состав матрикса. Это говорит о том, что контроль топологии ДНК — важная составляющая функции обеих структур.

28.7 Хроматин подразделяют на эухроматин и гетерохроматин

Ключевые понятия:

- **Обособленные хромосомы можно наблюдать только при митозе**
- **Во время интерфазы хроматин, по большей части, пребывает в форме эухроматина, который упакован гораздо менее плотно, чем митотические хромосомы.**
- **Участки конститутивного гетерохроматина остаются плотно упакованными в продолжение всей интерфазы.**

Каждая хромосома содержит одну очень длинную молекулу ДНК. Эта двухцепочечная молекула уложена в фибриллу, которая непрерывна на всем пространстве хромосомы. *Поэтому, говоря о структурах интерфазного хроматина и метафазных хромосом, необходимо указать, каким образом одна неимоверно*



РИС. 28.11 Митотическая пара сестринских хроматид состоит из компактно сложенного волокна диаметром ~30 нм. Фотография любезно предоставлена Daniel L. Hartl, Harvard University

длинная молекула ДНК оказывается упакованной так, что может быть не только транскрибирована и реплицирована, но и подвергнута циклическим переходам между большей и меньшей плотностями упаковки.

Хромосомы эукариот обретают морфологическую индивидуальность только на короткий промежуток времени, когда клетка делится. Только тогда каждую из них можно увидеть как компактную единицу. На РИС. 28.11 представлена электронная микрофотография пары сестринских хроматид на стадии метафазы. (Сестринские хроматиды — это дочерние хромосомы, продукт предшествующего митозу события репликации; в метафазе митоза они все еще соединены вместе.) Каждая хроматида состоит из волокна диаметром ~30 нм, узловатого на вид. ДНК хромосом в 5–10 раз более конденсирована (плотно упакована), чем ДНК интерфазного хроматина.

И все же в продолжение большей части жизненного цикла эукариотической клетки ее генетический материал визуально разбросан по ядру, и в нем нельзя различить отдельные хромосомы. От митоза до митоза структура интерфазного хроматина не претерпевает никаких видимых изменений, в частности, никаких возмущений в период репликации, когда количество хроматина удваивается. Хроматин на вид волокнистый (фибриллярный), но разобраться в деталях конфигурации волокна в пространстве нелегко. По крайней мере, само волокно (фибрилла) то же, что и в митотических хромосомах.

Микрофотография среза через область ядра на РИС. 28.12 демонстрирует, что хроматин может быть поделен на два типа:

- Большинство участков хроматина упакованы гораздо менее плотно, чем в митотической хромосоме. Такой материал называют **эухроматином**. Он выглядит рассеянным по ядру и занимает большую часть ядерной области на рис. 28.12.
- Некоторые участки хроматина упакованы очень плотно — до состояния, сравнимого с митотической хромосомой. Такой материал называют **гетерохроматином**. В его состав входят ДНК центромер и некоторые другие участки генома, которые за все время клеточного цикла претерпевают относительно малые колебания конденсированности. На рис. 28.12 гетерохроматин представлен серией обособленных комков. Нередко разнообразные участки гетерохроматина слипаются в густо окрашиваемый **хромоцентр** (это описание относится к участкам, которые гетерохроматичны всегда, что называется, конститутивно; кроме конститутивного гетерохроматина есть еще факультативный, который представляет собой участки эухроматина, временно перешедшие в конденсированное состояние).

Пространственно, эухроматин и гетерохроматин непрерывно переходят один в другой, и соединения имеют вид того же волокна. Судя по всему, два типа хроматина отображают разные степени конденсированности ДНК. Выходит, что генетический материал организован в манере, допускающей не только поддержание его в альтернативных состояниях эухроматина и гетерохроматина на участках, расположенных бок о бок, но и циклические изменения плотности упаковки эухроматина сообразно интерфазам и делениям. Молекулярная природа этих состояний обсуждается в гл. 30 «Контроль структуры хроматина».



РИС. 28.12 На тонком срезе ядра, окрашенного по Фельгену, гетерохроматин выглядит как скопление компактных затемнений у ядрышка и ядерной мембраны. Фотография любезно предоставлена Edmund Puvion, Centre National de la Recherche Scientifique

Структурное состояние генетического материала коррелирует с его активностью. Гетерохроматин отличается от эухроматина следующими особенностями:

- Он перманентно конденсирован.
- Его ДНК зачастую состоит из многочисленных повторов небольшого числа нетранскрибируемых последовательностей.
- Плотность расположения генов на участках конститутивного гетерохроматина, по сравнению с эухроматином, очень сильно снижена. Гены, при транслокации попавшие в гетерохроматин или его окрестности, часто становятся неактивными.
- ДНК конститутивного гетерохроматина реплицируется в поздней S-фазе; генетическая рекомбинация имеет в ней место с пониженной частотой — возможно, из-за ее повышенной конденсированности.

Для отслеживания изменения свойств ДНК и белковых компонентов хроматина используется несколько молекулярных индикаторов (см. разд. 31.3 «Существование гетерохроматина обусловлено взаимодействиями гистонов с другими белками»; например, пониженная ацетилированность гистонов, повышенная метилированность одного из них, гиперметилированность цитидиновых оснований в ДНК). Эти изменения на молекулярном уровне приводят к конденсации генетического материала, которая обуславливает его инактивацию.

Активные гены находятся в эухроматине, но в каждый момент времени лишь малая часть всех последовательностей эухроматина подлежит транскрипции. Расположение гена в эухроматине *необходимо*, но само по себе *не достаточно* для его экспрессии.

28.8 У хромосом проявляются полосатые паттерны окрашивания

Ключевые понятия:

- **Определенные методики окрашивания придают хромосомам исчерченность в виде серий G-полос, или Гимза-дисков.**
- **В полосах содержание G-C ниже, чем в промежутках между ними.**
- **Гены сосредоточены в G-C-богатых светлых промежутках.**

Из-за диффузного состояния хроматина изучать специфичность его организации прямыми методами невозможно. Более перспективный ход — задаться вопросом: упорядочена ли структура митотических хромосом? Зафиксировано ли соответствие конкретной геномной последовательности ее местоположению на хромосоме, или, может быть, укладка волокна — не такое уж упорядоченное событие?

Каждая хромосома обладает своеобразной воспроизводимой ультраструктурой. Если фиксированные и определенным образом обработанные хромосомы окрасить по Гимза, в них проявятся серии **G-полос**. Окрашенный препарат диплоидного набора хромосом человека показан на рис. 28.13.

После того как была разработана эта методика, хромосомы научились различать не только по размерам и расположениям центромер. Маркировка в виде G-полос позволяет идентифицировать хромосомы по характерным паттернам исчерченности. Такой паттерн служит разметкой, с помощью которой можно устанавливать факты транслокаций материала между хромосомами, пользуясь в качестве эталона окрашенным препаратом исходного диплоидного набора. На рис. 28.14 схематично изображена X-хромосома человека. Каждая полоса соответствует крупному отделу генома, ~10^7 пн, где могут находиться сотни генов.

Методика окрашивания хромосом имеет громадное практическое значение, но механизм возникновения исчерченности пока остается нераскрытым. Известно только, что необработанные хромосомы окрашиваются более или менее однородно. Это означает, что проявление паттерна полос зависит от предварительной обработки. Способов такой обработки известно несколько. Предположительно, предварительная обработка экстрагирует из промежутков между полосами компонент, который связывает краску. Из-за этого само окрашивание оставляет промежутки бледными.

Полосы отличаются от промежутков более низким содержанием G-C. Это единственное известное отличие, причем очень интересное. Если в большой хромосоме, содержащей 10^8 пн ДНК (100 Мпн), насчитывается 10 полос, это означает, что она состоит из участков по 5×10^6 пн ДНК (5 Мпн), чередующихся по G-C-составу. Как показывают результаты гибридизации хромосом с мРНК, большинство генов расположено не в полосах, а в промежутках. Все это говорит о подчиненности генетического материала каким-то дальнодействующим факторам организации, обуслов-



Рис. 28.13 На специально окрашенных препаратах каждую хромосому гаплоидного набора можно идентифицировать по характерному паттерну исчерченности. Фотография любезно предоставлена Lisa Shaffer, Washington State University–Spokane

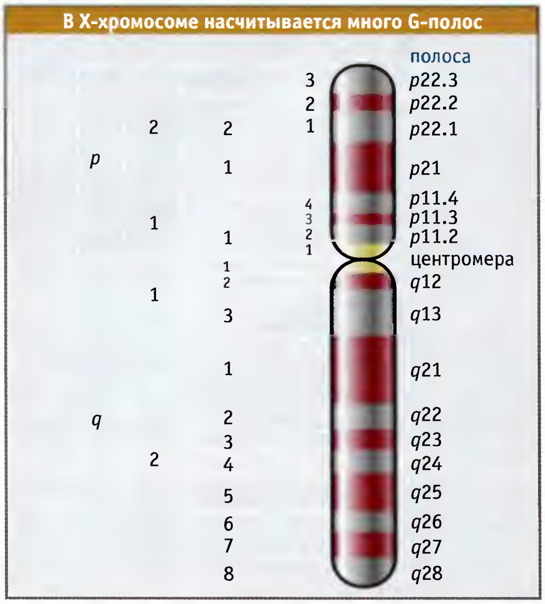


РИС. 28.14 В X-хромосоме человека, окрашенной по Гимза, четко различим характерный паттерн поперечной исчерченности. Короткое плечо хромосомы обозначают как *p*, а длинное – как *q*. Каждое плечо хромосомы подразделяют на крупные области (полосы и промежутки), которые, в свою очередь, подразделяются на более мелкие области. Данная карта имеет низкое разрешение. При более высоком разрешении некоторые из полос не выглядят цельными: например, полоса *p*21 делится на *p*21.1, *p*21.2 и *p*21.3

ленным, однако, на уровне нуклеотидных последовательностей.

Эти наблюдения подтверждаются анализом полной нуклеотидной последовательности человеческого генома. Из РИС. 28.15 очевидно существование отчетливых флуктуаций G-C-состава в геноме, поделенном на небольшие транши (сегменты, отрезки ДНК). Средний

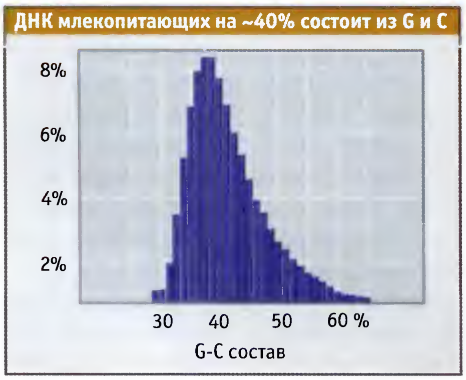


РИС. 28.15 На коротких отрезках генома G-C состав может сильно колебаться. Каждый столбик диаграммы соответствует процентному содержанию 20 тпн фрагментов с конкретным значением G-C-состава

показатель в 41% G-C характерен для геномов млекопитающих. На некоторых участках содержание G-C может достигать 65% или быть пониженным до 30%. Если длину анализируемых траншей увеличить, разброс по G-C-составу снизится. Средняя длина участка с >43% G-C составляет 200–250 тпн. Из этого ясно, что чередование полос и промежутков *не* отображает чередования гомогенных участков, различающихся по G-C-составу, хотя полосы действительно содержат большее число G-C-бедных сегментов, чем промежутки. Гены сконцентрированы на участках повышенного содержания G-C. Каким образом распределение содержания G-C влияет на структуру хромосомы, еще предстоит понять.

28.9 Хромосомы типа ламповых щеток имеют выпетливания

Ключевые понятия:

- У хромосом типа ламповых щеток сайты экспрессии генов выглядят как петли, выпущенные из оси хромосомы.

Исключительно полезно было бы воочию пронаблюдать за экспрессией генов: структурные изменения, сопутствующие транскрипции, стали бы явными. Но из-за уплотненности ДНК в составе хроматина, в сочетании со сложностью идентификации в ней конкретных генов, визуализацию транскрипции индивидуальных активных генов приходится признать невозможной.

И все же в жизни некоторых клеток бывают необычные ситуации, когда экспрессию генов можно *увидеть*. Иногда хромосомы можно встретить в форме, деконденсированной настолько, что можно различить отдельные локусы (или группы локусов). Латеральная дифференциация структуры очевидна у многих хромосом при первом появлении в мейозе. На этой стадии они напоминают нити бус. «Бусины» — густо окрашиваемые гранулы — правильно называть **хромомерами**. Во время мейоза гены экспрессируются очень слабо, и на практике, для идентификации активностей индивидуальных генов, такой материал бесполезен. Исключение, подходящее для такого рода исследований, представляют **хромосомы типа ламповых щеток**, лучше всего изученные у некоторых амфибий.

Хромосомы типа ламповых щеток образуются во время необычайно продолжительного мейоза, который может длиться до нескольких месяцев! В продолжение этого периода хромосомы удерживаются в распущенной форме, что можно наблюдать в световой микроскоп. На более поздней стадии мейоза хромосомы возвращаются к своим обычным компактным размерам.

Хромосомы типа ламповых щеток — это мейотические биваленты, каждый из которых состоит из двух



РИС. 28.16 Хромосома типа ламповой щетки представляет собой мейотический бивалент: две пары сестринских хроматид удерживаются вместе из-за хиазм (обозначенных стрелками). Фотография любезно предоставлена Joseph G. Gall, Carnegie Institution

пар сестринских хроматид. На РИС. 28.16 приведен пример, в котором пары сестринских хроматид, в основном, разделились, и удерживаются вместе лишь хиазмами. Каждая пара сестринских хроматид образует серию эллипсоидальных хромомеров, ~1–2 мкм в диаметре, соединенных очень тонкой нитью. Эта нить содержит два сестринских дуплекса ДНК и непрерывно пробегает всю хромосому, через хромомеры.

Длины хромосом типа ламповых щеток у тритона *Notophthalmus viridescens* колеблются от 400 до 800 мкм, по сравнению с 15–20 мкм, наблюдаемыми у них же позднее в мейозе. Это означает, что хромосомы типа ламповых щеток упакованы в ~30 раз более рыхло, чем обычные мейотические. Суммарная длина набора хромосом типа ламповых щеток равняется 5–6 мм и распределена по ~5000 хромомеров.



РИС. 28.17 Петля хромосомы типа ламповой щетки окутана рибонуклеопротеиновым матриксом. Воспроизведено по Gall, J. G., et al. Mol. Biol. Cell. December, 1999. 10: 4385–4402. Фотография любезно предоставлена Joseph G. Gall, Carnegie Institution

Название «хромосомы типа ламповых щеток» взялось из описания латеральных петель, которые выпячиваются из хромомеров по определенным положениям. (Эти петли делают всю хромосому похожей на ламповую щетку, ныне исчезнувший предмет обихода.) Петли выпячиваются парами, по одной от каждой сестринской хроматиды и непрерывно продолжают аксиальную нить хромосомы. Они — не что иное, как материал хромосомы, выпяченный из хромомера, в котором он был организован более компактно.

Петли окружены матриксом из рибонуклеопротеинов (РНП), где содержатся новосинтезированные цепи РНК. Нередко, по увеличению длины РНП, огибающих петлю, можно вычленить единицу транскрипции. Пример приведен на РИС. 28.17.

Таким образом, в виде петли из хромосомы выпущен сегмент ДНК, который активно транскрибируется. В ряде случаев были идентифицированы петли, соответствовавшие конкретным генам. В такой ситуации структура транскрибируемого гена — и все, что касается его продукта — может стать предметом изучения *in situ*.

28.10 Политенные хромосомы выглядят полосатыми

Ключевые понятия:

- **Окрашенные препараты политенных хромосом двукрылых насекомых несут серии полос, которые можно использовать как цитологическую карту.**

Интерфазные ядра в некоторых тканях личинок двукрылых мух содержат хромосомы, длина и диаметр которых, по сравнению с обычным состоянием, намного увеличены. В качестве примера на РИС. 28.18 приведен препарат набора хромосом из слюнной железы личинки *D. melanogaster*. Их называют **политенными** хромосомами.

Каждая хромосома политенного набора состоит из видимой серии **полос** (более правильно, но редко, адресуемым как хромомеры). Полосы различны по ширине, которая может составлять от 0,05 до 0,5 мкм. Самую узкую полосу можно различить только в электронный микроскоп. Полосы (в них содержится бо́льшая часть всей массы ДНК) интенсивно окрашиваются соответствующими реактивами. Участки между полосами окрашиваются в более светлый тон; их называют **промежутками**. В наборе хромосом *D. melanogaster* насчитывается ~5000 полос.

Центромеры всех четырех хромосом *D. melanogaster* сходятся вместе и образуют хромоцентр, состоящий по большей части из гетерохроматина. (У самцов в его состав входит вся Y-хромосома.) За вычетом хромоцентра получается, что 75% геномной ДНК *D. melanogaster* организовано в виде чередующихся полос и промежутков. Суммарная длина хромосом в гаплоидном наборе равняется 2000 мкм, если содержащуюся в них ДНК полностью



РИС. 28.18 Политенные хромосомы *D. melanogaster* представляют собой серии чередующихся полос и промежутков. Фотография любезно предоставлена José Bonner, Indiana University

расправить, она растянется на 40000 мкм. Соотношение длин дает коэффициент упаковки ~20, что наглядно поясняет, в сколь деконденсированном состоянии пребывает генетический материал политенных хромосом по сравнению с интерфазным хроматином и митотическими хромосомами (у которых коэффициенты упаковки равняются ~10^3 и ~10^4 соответственно).

Как устроены эти гигантские хромосомы? Каждая из них — продукт последовательных репликаций синапсированной диплоидной пары хромосом. В своем расправленном состоянии реплики не разделяются, а остаются скрепленными. Предположим, на старте процесса содержание ДНК в синапсированной паре хромосом равняется 2С, где С –это количество ДНК в одной хромосоме. Затем содержание ДНК удваивается, затем — снова, и так до девяти раз подряд. Как максимум, содержание ДНК при этом повышается до 1024С. Число удвоений зависит от конкретной ткани личинки *D. melanogaster*, где происходит процесс.

На препаратах политенные хромосомы выглядят состоящими из большого числа продольно идущих параллельных волокон, которые плотно конденсированы в зонах полос и более рыхло — в промежутках. Судя по всему, каждое волокно (фибрилла) представляет собой одинарную гаплоидную хромосому (с содержанием ДНК, равным С). Степень политении — это число одинарных хромосом в составе гигантской хромосомы.

У каждой линии *Drosophila* — свой паттерн исчерченности хромосом. Постоянное число и линейное расположение полос впервые заметили в 1930-х годах и сразу поняли, что они образуют *цитологическую карту* хромосом. Хромосомные перестройки — делеции, инверсии, дупликации — приводят к изменениям в порядке расположения полос.



РИС. 28.19 При помощи гибридизации *in situ* в политенных хромосомах можно идентифицировать полосы, содержащие конкретные гены

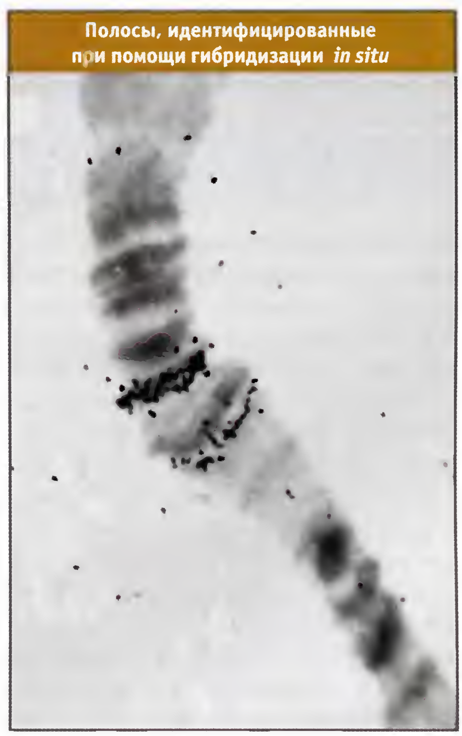


РИС. 28.20 Полосы 87А и 87С: при большом увеличении видно, что их ДНК специфично гибридизуется с РНК, которая была выделена из клеток, подвергавшихся тепловому шоку. Фотография любезно предоставлена José Bonner, Indiana University

Линейный массив полос можно рассматривать как отображение линейного массива генов. Генетические перестройки, нанесенные на карту сцепления признаков могут быть соотнесены с изменениями в цитологической карте. В идеале, конкретная мутация может быть запеленгована в конкретной полосе. Суммарное чило генов *D. melanogaster* превышает число полос, и в большинстве полос (если не во всех) содержится по нескольку генов.

Позиции, занимаемые конкретными генами на цитологической карте, можно определить непосредственно. Для этого используются методики **гибридизации *in situ***; их обобщенный протокол приведен на РИС. 28.19. Радиоактивная проба, представляющая ген (чаще всего меченая кДНК) гибридизуется *in situ* с денатурированной ДНК политенных хромосом. Позиции генов, соответствующих пробе, определяют авторадиографией, по оптическому наложению засвеченных зерен фотоэмульсии на конкретную полосу (или полосы). Пример показан на РИС. 28.20. В новом веке на смену радиоактивным пробам пришли флуоресцентные. Новейший тип методик с их использованием позволяет напрямую локализовать конкретную последовательность в конкретной полосе.

28.11 Политенные хромосомы образуют «вздутия» в местах экспрессии генов

Ключевые понятия:

- На окрашенных препаратах политенных хромосом полосы, где шла активная экспрессия генов, выглядят раздутыми (имеют вид «пуфов»).

Политенные хромосомы обладают той интригующей особенностью, что их транскрипционно-активные области можно увидеть воочию. Участки хромосом, соответствующие отдельным полосам, на какой-то промежуток времени переходят в расплетенное состояние. На гистологических препаратах такая полоса выглядит как **пуф** (вздутие), материал которого как бы выдавлен из хромосомы перпендикулярно ее оси. Особо крупные пуфы (их называют кольцами Бальбиани) можно видеть на РИС. 28.21. Как же они устроены?

В месте пуфа волокна политенной хромосомы перешли в расплетенное состояние — по сравнению с обычным, более плотно упакованным состоянием, имеющим вид полосы. При этом сами волокна сохраняют непрерывность и составляют одно целое с волокнами осевой части хромосомы. Обычно пуфы — это производные отдельных полос; если они очень большие, как кольца Бальбиани, то вздутие может закрывать обзор соседних полос.

Паттерн, образуемый пуфами на хромосомах, может быть соотнесен с профилем экспрессии генов. В процессе личиночного развития *Drosophila* возникновение и регресс пуфов подчинены определенным тканеспецифичным закономерностям. Для каждой ткани в каждый момент времени есть свой характерный паттерн пуфов. Возникновение пуфов индуцируется гормоном экдизоном, контролирующим развитие *Drosophila*. Некоторые пуфы индуцируются экдизоном непосредственно, другие — косвенно, через продукты более ранних пуфов.

Пуф — это *место, где идет синтез РНК*. Традиционное видение таково, что пуфы возникают как результат релаксации хромосомы, необходимой для синтеза РНК. Пуф может отражать активное состояние одного единственного гена. «Вздутые» полосы отличаются от обыкновенных полос тем, что в них накапливаются дополнительные белки, в числе которых — РНК-полимераза II и другие белки, ассоциируемые с транскрипцией.

Свойства политенных хромосом и хромосом типа ламповых щеток наводят на мысль: для того чтобы быть транскрибированным, генетическому материалу приходится выходить из своего обычного плотно упакованного состояния. Все же следует иметь в виду, что вопрос, насколько хорошо локальное рассредоточение материала в огромной хромосоме имитирует молекулярные события в интерфазном эухроматине, остается открытым.

Имеет ли исчерченность политенных хромосом какой-то функциональный смысл, то есть, соответствует ли каждая полоса какой-то генетической единице? Естественно было бы предположить, что, имея в распоряжении полностью отсеквенированный геном мухи, ответить на этот вопрос довольно легко. В самом деле, прокартировав промежутки между полосами на полноразмерной пос-



РИС. 28.21 В клетках слюнных желез насекомого *C. tentans* хромосома IV имеет три кольца Бальбиани. Воспроизведено по *Cell*, vol. 4, Daneholt, B., et al., Transcription in polytene chromosomes . . ., pp. 1–9. Copyright 1975, с разрешения Elsevier. Фотография любезно предоставлена Bertil Daneholt, Medical Nobel Institute

ледовательности генома, можно, наверное, определить, есть ли у полос конкретная единообразная сущность. На сегодняшний день никакого паттерна геномных последовательностей, который указывал бы на функциональную значимость полос, обнаружено не было.

28.12 Эукариотические хромосомы служат клетке приспособлениями для сегрегации генетического материала

Ключевые понятия:

- Эукариотические хромосомы удерживаются в веретене деления благодаря микротрубочкам, прикрепленным к кинетохорам, формирующимся при центромерах.
- Во многих случаях центромеры содержат гетерохроматин, богатый последовательностями сателлитной ДНК.

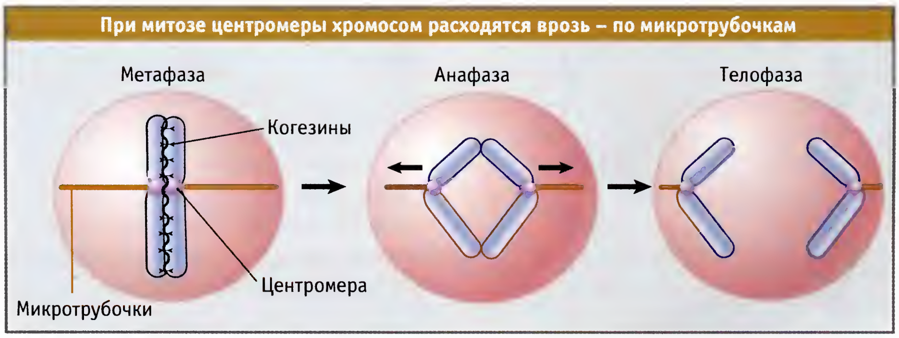
Во время митоза сестринские хроматиды расходятся к противоположным полюсам клетки. Их движение зависит от прикрепления хромосомы к микротрубочкам, которые другим своим концом сообщаются с полюсами клетки. (Система микротрубочек — это компонент цитоскелета; при митозе микротрубочки реорганизуются и соединяют хромосомы с полюсами клетки.) Цитоплазматические сайты, где удерживаются концы микротрубочек, называют центрами организации микротрубочек (**ЦОМ**). Они расположены близко к центриолям на полюсах, а также на хромосомах.

Рис. 28.12 иллюстрирует разделение сестринских хроматид по мере того, как митоз прогрессирует от метафазы к телофазе. Участок хромосомы, ответственный за ее сегрегацию (распределение по дочерним клеткам) при митозе и мейозе, называют **центромерой**. Микротрубочки тянут центромеры сестринских хроматид к противоположным полюсам. В противовес этой движущей силе, белки когезины, которыми «склеены» сестринские хроматиды, продолжают удерживать их вместе. Разделение сестринских хроматид начинается с центромер, а затем они полностью освобождаются друг от друга (поскольку когезины деградируют). Центромеры растягиваются к полюсам, увлекая за собой хромосомы (бывшие сестринские хроматиды). Таким образом, хромосома работает как связующее приспособление между большим числом содержащихся в ней генов и аппаратом клеточного деления. Хромосома (адаптор) координирует распределение генетического материала с фактическим пространственным разделением вещества клетки.

У хромосомы есть область, через которую сестринские хроматиды удерживаются вместе до того, как отделятся друг от друга и станут хромосомами. На фото рис. 28.11, где показаны сестринские хроматиды на момент метафазы митоза, эта область выглядит как продольная перетяжка, с которой соприкасаются все четыре плеча хроматид.

Центромера принципиально важна для сегрегации, об этом свидетельствует поведение хромосом, потерпевших разломы. После разлома по одному из плеч, дистальный обломок этого плеча, отделившись от остальной хромосомы, образует **ацентрический фрагмент**. Из-за отсутствия центромеры ацентрический фрагмент не может быть прикреплен к веретену деления, и в результате он не попадает в состав ни одного из дочерних ядер.

В своем движении хромосомы, в качестве навигаторов, полагаются на дискретные центромеры, и центромера в хромосоме может быть *только* одна. Когда, в результате транслокаций, образуются хромосомы с более чем одной центромерой, митоз превращает их в аберрантные структуры. Это происходит из-за того, что две центромеры, принадлежащие *одной и той же* сестринской хроматиде, могут быть растянуты к разным полюсам, и хромосома при этом рвется. У некоторых видов



В анафазе митоза микротрубочки, прикрепленные к центромерам, тянут хромосомы к противоположным полюсам клетки. До наступления анафазы сестринские хроматиды удерживаются вместе белками когезина. На иллюстрации центромера изображена посередине хромосомы (метацентрической). На самом деле центромера может находиться в любом месте на оси хромосомы, например, рядом с одним из концов (у акроцентрических хромосом) или на самом конце (у телоцентрических хромосом)



РИС. 28.23 Избирательное интенсивное окрашивание центромер во всех хромосомах можно получить с помощью гистологической техники C-banding. Фотография любезно предоставлена Lisa Shaffer, Washington State University–Spokane

организмов центромеры «диффузные» и ведут себя совсем иным образом. Пока только дискретные центромеры были проанализированы на молекулярном уровне.

Во многих случаях участки хромосом по бокам от центромеры содержат много сателлитной ДНК и, в основном, соответствуют гетерохроматину. Во время митоза хромосома конденсирована; центромерный гетерохроматин в ней не различим. Визуализировать центромерный гетерохроматин можно при помощи методики окрашивания, проявляющей в хромосомах **С-полосы**. На фотографии препарата, приведенной на РИС. 28.23, все центромеры видны как густо окрашенные участки. Хотя так обычно и бывает, все же выявить гетерохроматиновую окрестность можно не у каждой центромеры, и, возможно, этот гетерохроматин не так уж существен для сегрегации.

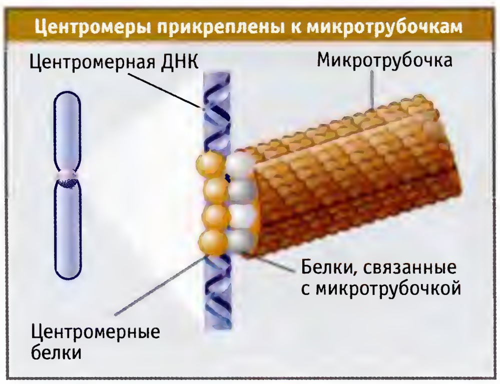


РИС. 28.24 Центромера соответствует элементу хромосомной ДНК со специфической последовательностью: этот элемент связывает определенные белки, оформляюшие центромеру как место на хромосоме, с которым могут взаимодействовать другие белки – связанные с микротрубочками

Участок хромосомы обретает свойства центромеры в силу наличия в нем определенных последовательностей ДНК (хотя сами последовательности определены у довольно узкого круга объектов). Центромерная ДНК связывает специфические белки, порождающие структуру, которая прикрепляет хромосому к микотрубочкам. Эта структура — **кинетохор** — выглядит как темноокрашенный волокнистый объект размером ~400 нм. Кинетохор образует на хромосоме центр организации микротрубочек. На РИС. 28.24 изображена схема иерархически организованной структуры, осуществляющей коммуникацию между центромерной ДНК и микротрубочкой. Белки, связанные с ДНК центромеры, связывают другие белки, которые, в свою очередь, связываются с микротрубочкой (см. разд. 28.15 «Центромера связывается с белковым комплексом»).

28.13 Центромеры содержат повторяющиеся последовательности ДНК

Ключевые понятия:

- Центромеры в хромосомах высших эукариот содержат большие количества повторов ДНК.
- Функция этих повторов пока не известна.

Протяженность участка, выступающего в качестве центромеры, у большинства хромосом довольно велика. (Короткие дискретные элементы, играющие роль центромер у *S. cerevisiae* — это, скорее, исключение.) В тех случаях, когда центромерную ДНК можно отождествить со специфическими последовательностями, эти последовательности обычно содержат повторяющиеся нуклеотидные мотивы.

Дрожжи *S. cerevisiae* — это, пока что, единственный случай, для которого центромерная ДНК была идентифицирована по своей способности придавать стабильность плазмидам. Этот подход применяли и к другому виду дрожжей, *Schizosaccharomyces pombe*. У *S. pombe* всего три хромосомы; в каждой из них удаляли большинство последовательностей и отслеживали стабильность получившейся минихромосомы. Таким методом пытались идентифицировать минимальную геномную область, способную функционировать в качестве центромерной ДНК. Оказалось, что центромерная ДНК *S. pombe* содержится в пределах участков по 40–100 пн (в зависимости от конкретной хромосомы), и что эти участки состоят (всецело или по большей части) из повторов ДНК. Какая доля от каждого из этих, довольно протяженных, участков требуется для сегрегации хромосом при митозе и мейозе, пока не ясно.

Попытки локализовать центромерную ДНК в геноме *Drosophila* показали, что ее последовательности рассеяны на больших участках (от 200 до 600 тпн). Значительная протяженность такого рода центромерных областей

генома позволяет предполагать наличие в них нескольких функционально обособленных отделов (например, последовательностей, необходимых для сборки кинетохора или попарного скрепления сестринских хроматид).

Размеры центромерных областей в геномах мухи *Drosophila* и растения *Arabidopsis* сравнимы друг с другом. ДНК каждой из пяти хромосом *Arabidopsis* содержит область центромеры, где рекомбинация очень сильно подавлена. Эта область занимает >500 тпн. В нее совершенно точно укладывается вся центромерная ДНК, но прямых сведений о том, какая доля этой области требуется для нормальной функции, пока нет. В данной области есть экспрессируемые гены, что вселяет некоторые сомнения в центромерном предназначении всей области. Середину области занимают повторы по 180 пн; это как раз тип структуры, в общем случае ассоциируемый с центромерами. В данном случае говорить о доминирующем функциональном значении именно этих повторов преждевременно.

В гетерохроматине центромер приматов преобладает α-сателлитная ДНК, представляющая собой тандемный массив повторяющихся единиц по 170 пн. Различия между отдельными повторами довольно значительны, хотя те α-повторы, что находятся в центромере, похожи друг на друга сильнее, чем на повторы того же семейства из других областей хромосомы. Ясно, что последовательности, обеспечивающие функциональность центромеры, находятся где-то среди блоков α-сателлитной ДНК; сами ли это блоки α-повторов или же другие геномные последовательности, находящиеся в их окружении, неизвестно.

28.14 Центромеры *S. cerevisiae* несут короткие последовательности ДНК

Ключевые понятия:

- Последовательности *CEN* у *S.cerevisiae* были идентифицированы как придающие плазмиде способность к правильной митотической сегрегации.
- Последовательности *CEN* состоят из коротких консервативных элементов *CDE-I* и *CDE-III*, которые фланкируют А-Т-богатый элемент *CDE-II*.

Если центромерная последовательность ДНК в самом деле отвечает за сегрегацию, то любая молекула ДНК, несущая такую последовательность, должна при клеточном делении двигаться правильным образом, тогда как ДНК, не имеющая в своем составе такой последовательности, должна испытывать проблемы с сегрегацией. Именно таким соображением руководствовались, чтобы вычленить центромерные последовательности ДНК у дрожжей *S. cerevisiae*. Хромосомы дрожжей не имеют выраженных кинетохоров, присущих высшим эукариотам, но во всех остальных отношениях их сегрегация при митозе и мейозе протекает по тем же самым механизмам.

Методами генной инженерии изготавливают дрожжевые плазмиды, реплицируемые наподобие хромосомных последовательностей (см. разд. 15.8 «Лицензирующие факторы держат под контролем повторную репликацию геномов эукариот»). Однако при митозе и мейозе такие плазмиды выказывают нестабильность и большинством клеток теряются, поскольку сегрегируют неупорядоченно. Центромерные участки хромосомной ДНК были идентифицированы как последовательности, придающие плазмидам стабильность.

Фрагмент *CEN* (centromeric DNA regions fragment; фрагмент участков центромерной ДНК) определяется как минимальная область, которая может придать реплицируемой плазмиде митотическую стабильность.

Фрагментом *CEN* одной из хромосом, можно, безо всяких явных последствий, заменить центромеру другой хромосомы. Это говорит о взаимозаменяемости центромер. *Они используются просто для прикрепления хромосом к веретену деления и не играют никакой роли в создании различий между хромосомами.*

Последовательности *CEN* умещаются на отрезке длиной 120 пн. Они упакованы в виде структуры, устойчивой к нуклеазам, и связывают (через белки) ровно одну микротрубочку. Чтобы идентифицировать белки, с которыми связана центромерная ДНК, а также белки, соединяющие хромосому с веретеном деления, следует ознакомиться с центромерой *S. cerevisiae* в деталях.

Как показано на РИС. 28.25, в составе *CEN* можно выделить три отдельных элемента (последовательности):

- Элемент *CDE-I* (cell cycle-dependent element; зависимый от клеточного цикла элемент) — это консервативная последовательность 9 пн (с минорными

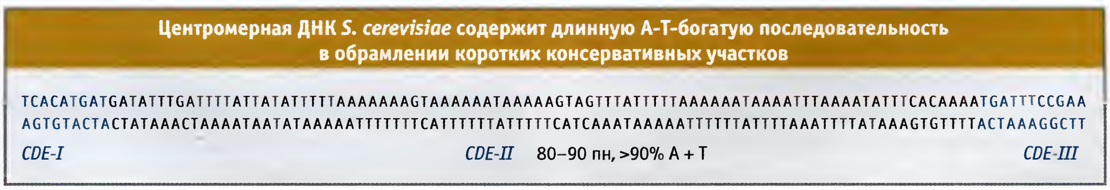


РИС. 28.25 Элемент *CEN* содержит длинную А-Т-богатую последовательность в обрамлении коротких консервативных участков. Этот вывод сделан на основании оценки сходства между частями нуклеотидных последовательностей центромерной ДНК у дрожжей

вариациями), расположенная по левому краю всех центромер.

- Элемент *CDE-II*, с А-Т-богатой последовательностью от 80 до 90 пн, обнаруживается во всех центромерах; его функциональность определяется скорее длиной, нежели точным нуклеотидным мотивом. Из-за специфического состава оснований (избытка А и Т) в *CDE-II* могут возникать характерные искажения двуспиральной структуры ДНК. Своим строением *CDE-II* напоминает некоторые сателлитные ДНК, состоящие из коротких тандемных повторов (см. разд. 6.12 «Сателлитная ДНК членистоногих состоит из очень коротких идентичных повторов»).
- Элемент *CDE-III*, высококонсервативная последовательность из 11 пн, расположенная на правом краю всех центромер. По обе стороны от этого элемента находятся менее консервативные последовательности, но, скорее всего, тоже существенные для функциональности центромеры. (*CDE-III* может оказаться длиннее, чем 11 пн, если фланкирующие последовательности действительно существенны и составляют с элементом одно функциональное целое.)

Мутации в CDE-I или CDE-II снижают функциональность центромеры, но не инактивируют ее. Точечные мутации по центральному CCG в CDE-III, напротив, полностью инактивируют центромеру.

28.15 Центромера связывается с белковым комплексом

Ключевые понятия:

- На *CDE-II* формируется специализированный белковый комплекс, альтернативный обычной структуре хроматина.
- С *CDE-III* связывается белковый комплекс CBF3.
- Белки, соединяющие эти два комплекса, могут обеспечивать их связь с микротрубочками.

Белки, необходимые для функционирования последовательностей *CEN* как организаторов центромер, удалось идентифицировать Есть несколько генов, мутации в которых нарушают сегрегацию хромосом и чьи белковые продукты локализуются в центромерах. Вклад этих белков в структуру центромеры обобщает РИС. 28.26.

Связывание элемента *CDE-II* с белком Cse4 приводит к построению специализированной структуры хроматина. Белок Cse4 похож на один из гистонов, составляющих базовые субъединицы хроматина (см. разд. 31.3 «Существование гетерохроматина обусловлено взаимодействиями гистонов с другими белками»). Белок Mif2, возможно также входит в комплекс Cse4 с ДНК; по крайней мере, он с ним связан. Аналоги Cse4 и Mif2 есть в центромерах высших эукариот (CENP-A и CENP-C). Консервативность структуры говорит о том, что данное взаимодействие может быть универсальным аспектом устройства центромеры. Основу для взаимодействия создает сгибание ДНК *CDE-II* вокруг белкового агрегата; вероятно, помогает то, что А-Т-богатая ДНК *CDE-II* изогнута сама по себе.

Участки элемента CEN связывают конкретные белки

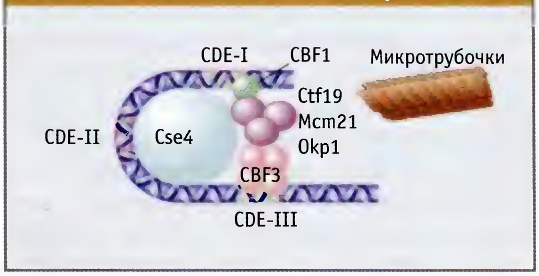


РИС. 28.26 На участке *CDE-II* двойная спираль ДНК обернута вокруг белкового агрегата, в котором есть Cse4p. Участок *CDE-III* взаимодействует с CBF-III, а участок *CDE-I* – с CBF1. Эти белки, в свою очередь, сообщаются друг с другом через Ctf19, Mcm21 и Okp1

С *CDE-I* связывается гомодимер CBF1; для функциональности центромер это взаимодействие не принципиально, но в его отсутствие точность сегрегации снижается примерно в 10 раз. С *CDE-III* связывается комплекс из четырех белков общей массой 240 кДа, называемый CBF3; для функционирования центромер это взаимодействие принципиально.

Белки, связанные с ДНК в *CDE-I* и *CDE-III*, физически взаимодействуют друг с другом, а также взаимодействуют с белковой структурой, привязанной к *CDE-II* через еще одну группу белков (Ctf19, Mcm21, Okp1). Соединение центромеры с микротрубочкой осуществляется именно через этот комплекс.

Общепринятая модель такова: комплекс Ctf19-Mcm21-Okp1 удерживается на центромере белковой структурой, похожей на нуклеосому — обычный строительный блок, из каких состоит хроматин. То, что по этой квазинуклеосоме проходит изгиб ДНК, позволяет белкам, связавшим фланкирующие элементы, стать частями единого комплекса. Некоторые компоненты этого комплекса (возможно, не те, что непосредственно связывают ДНК) связывают центромеру с микротрубочкой. По тому же принципу, и с использованием родственных компонентов, устроены кинетохоры самых разных организмов.

28.16 Теломеры содержат простые повторяющиеся последовательности ДНК

Ключевые понятия:

- Теломера нужна для стабильности конца хромосомы
- Теломерная ДНК состоит из многократного простого повтора, одна из цепей которого имеет вид $C_{>1}(A/T)_{1-4}$.

Другой неотъемлимый элемент всех хромосом — это **теломера**, которой «опечатан» ее конец. Теломера,



РИС. 28.27 Типичная теломерная ДНК имеет простую повторяющуюся первичную структуру, где G-T-богатая цепь выступает за пределы комплементарной С-А-богатой цепи. Выступ (так называемый G-хвост) образуется за счет ограниченной деградации С-А-богатой цепи

должно быть, имеет особую структуру, поскольку аберрантные концы хромосом, порожденные разломами и лишенные теломер, очень экспансивны и стремятся вступить в реакцию с другими хромосомами, тогда как в норме концы хромосом вполне стабильны.

Для идентификации последовательностей теломерной ДНК можно применить два критерия:

- Эти последовательности должны быть расположены на краю линейной хромосомы.
- Эти последовательности обязаны придавать стабильность линейной молекуле ДНК, находящейся внутри клетки (там же, где хромосомы).

Поиски модельной системы для постановки экспериментов на молекулярном уровне снова привели к дрожжам. Все плазмиды, которые существуют в культурах дрожжевых клеток благодаря обладанию последовательностями *ARS* и *CEN*, — это кольцевые молекулы ДНК. Линейные плазмиды в дрожжах нестабильны (клетка их разрушает). Может ли подлинная последовательность теломерной ДНК придать стабильность линейной плазмиде? В соответствующих экспериментах фрагменты ДНК дрожжей, располагавшиеся в концевых областях хромосом, а также концевая область линейной внехромосомной рДНК инфузории *Tetrahymena*. Наличия любой из этих последовательностей оказалось достаточно, чтобы стабилизировать дрожжевую плазмиду в линейной форме.

Опыты по идентификации теломерной ДНК коснулись самых разных групп высших и низших эукариот. Похоже, что устройство теломер подчинено чуть ли не единому универсальному принципу: у организмов, очень далеких друг от друга эволюционно (например, у растений и людей), обнаруживались теломерные последовательности одного и того же типа. Каждая теломера состоит из длинной серии коротких, тандемно повторяющихся последовательностей. Повторов может быть от 100 до 1000, в зависимости от организма, о котором идет речь.

Первичная структура (последовательность) теломерных повторов ДНК соответствует общей формуле $C_n(A/T)_m$, где $n > 1$, а m может принимать целочисленные значения от 1 до 4. Типичный пример приведен на РИС. 28.27. Теломерная ДНК обладает необычным свойством: ее G-T-богатая цепь длиннее своего С-А-богатого комплемента и образует концевой одноцепочечный выступ из 14–16 нуклеотидов. Вероятно, этот выступ (так называемый G-хвост) образуется как результат специфической ограниченной деградации С-А-богатой цепи.

Догадки о принципе функционирования теломеры строились на основании некоторых необычных свойств, которыми обладают природные линейные молекулы ДНК. Так, в популяции трипаносом длина концов линейных молекул ДНК варьирует. Если проследить за клоном отдельно взятой клетки, то можно увидеть, что теломера прирастает на 7–10 пн (то есть, на 1–2 повтора) за поколение. Еще больший интерес представляет судьба теломер инфузории, интродуцированных в ДНК дрожжей. После репликации в дрожжевой клетке *концы теломерной области Tetrahymena наращиваются теломерными повторами дрожжей.*

Наращивание концов хромосом теломерными повторами на каждом цикле репликации может решить проблему репликации концов линейной ДНК (обсуждавшуюся в разд. 16.2 «Края линейной ДНК сложно реплицировать»). Добавление повторов синтезом *de novo* противодействует потере повторов из-за неспособности репликации дойти до конца линейной хромосомы. При этом ее удлинение и укорачивание находятся в динамическом равновесии.

Если теломеры постоянно удлиняются (и укорачиваются), их точная последовательность не имеет значения. Все, что требуется — это чтобы их концы опознавались как субстрат, пригодный для наращивания. С учетом этого соображения поведение теломер инфузорий в дрожжах становится объяснимым.

28.17 Теломеры опечатывают концы хромосом

Ключевые понятия:

- Белок TRF2 катализирует реакцию, в ходе которой 3′-концевой повтор G-хвоста замещает гомологичный повтор, вытеснив его из дуплекса в вышележащей области теломерной ДНК.

Теломеры хромосом, если их экстрагировать из клетки, не ведут себя как одноцепочечная ДНК; вместо этого они проявляют аберрантную электрофоретическую подвижность и другие свойства.

Гуаниновые азотистые основания обладают необычной способностью к ассоциации друг с другом. Одноцепочечный G-хвост ДНК в составе теломеры может формировать вторичную структуру, в которой присутствуют «квартеты» гуанинов. В каждом квартете содержится четыре гуанина, соединенных водородными связями и лежащих в одной плоскости. Каждый из гуанинов квартета берется с гомологичной позиции повторяющейся

гексануклеотидной единицы TTAGGG. На РИС. 28.28 изображена новейшая модель, основанная на недавнем анализе кристаллической структуры. Квартет, изображенный на этом рисунке, представляет собой ассоциацию первых гуанинов четырех тандемных повторов. Он лежит поверх другого квартета с такой же организацией, сформированного из вторых гуанинов тех же четырех повторов. Серия квартетов может образовывать спиралевидную стопку. Образование структуры, где раскрываются необычные стереохимические свойства G-богатой последовательности, *in vitro*, конечно же, не означает, что квартеты точно так же формируются *in vivo*.

Что именно в устройстве теломеры обеспечивает стабильность конца хромосомы? На РИС. 28.29 показано, что в теломере образуется петля ДНК. Отсутствие свободного конца ДНК, возможно, и есть обстоятельство, стабилизирующее конец хромосомы. Размер петли в клетках животных составляет, в среднем, от 5 до 10 тпн.

На РИС. 28.30 показано, как образуется петля. Свободный одноцепочечный 3'-конец теломерной ДНК, $(TTAGGG)_n$, замещает гомологичный участок в вышележащей двухцепочечной области теломерной ДНК. Результат замещения напоминает образование D-петли при гомологичной рекомбинации: вытесненная серия повторов TTAGGG становится одноцепочечным участком, а G-хвост теломеры, который произвел это вытеснение, взаимодействует с комплементарной цепью ДНК.

Реакцию катализирует белок TRF2, который вместе с другими белками образует комплекс, стабилизирующий конец хромосомы. Доказательством важности его участия в защите концов хромосом служит то, что делеция его гена приводит к хромосомным перестройкам.



РИС. 28.29 Край линейной хромосомной ДНК образует петлю. Фотография любезно предоставлена Jack Griffith, University of North Carolina at Chapel Hill

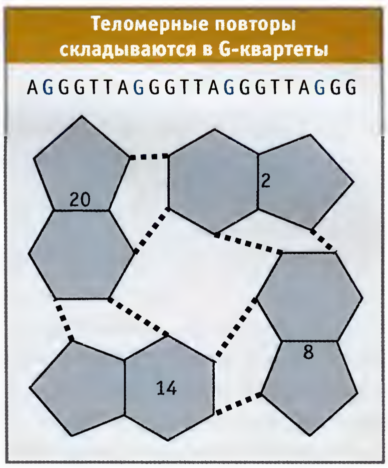


РИС. 28.28 Кристаллографический анализ структуры коротких теломерных повторов в одной из хромосом человека показал наличие стопки гуаниновых квартетов. Верхний из квартетов содержит первый гуанин каждого повтора. Под ним находится квартет вторых гуанинов (G3, G9, G15, G21), еще ниже – квартет третьих гуанинов (G4, G10, G16, G22)

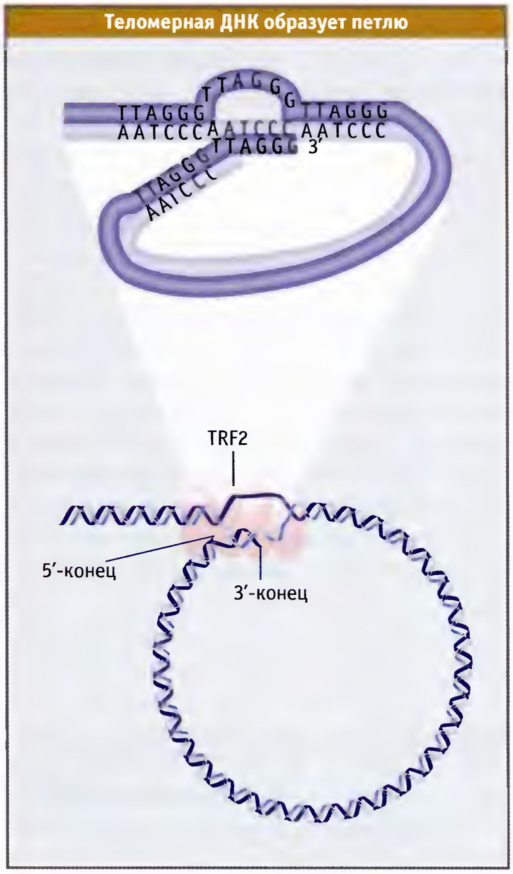


РИС. 28.30 3'-конец одиночной цепи теломерной ДНК $(TTAGGG)_n$ вытесняет из дуплекса гомологичную цепь. В результате образуется D-петля. Реакцию проводит белок TRF2

28.18 Теломеры синтезируются рибонуклеопротеиновым ферментом

Ключевые понятия:

- Теломераза использует 3′-ОН-группу G-T-участка теломеры как затравку для синтеза тандемных повторов TTGGGG.
- РНК-компонент теломеразы обладает последовательностью, идентичной С-А-богатым повторам теломерной ДНК.
- Каталитическая субъединица теломеразы — это обратная транскриптаза, использующая РНК-компонент как матрицу для синтеза G-T-богатой последовательности теломерной ДНК.

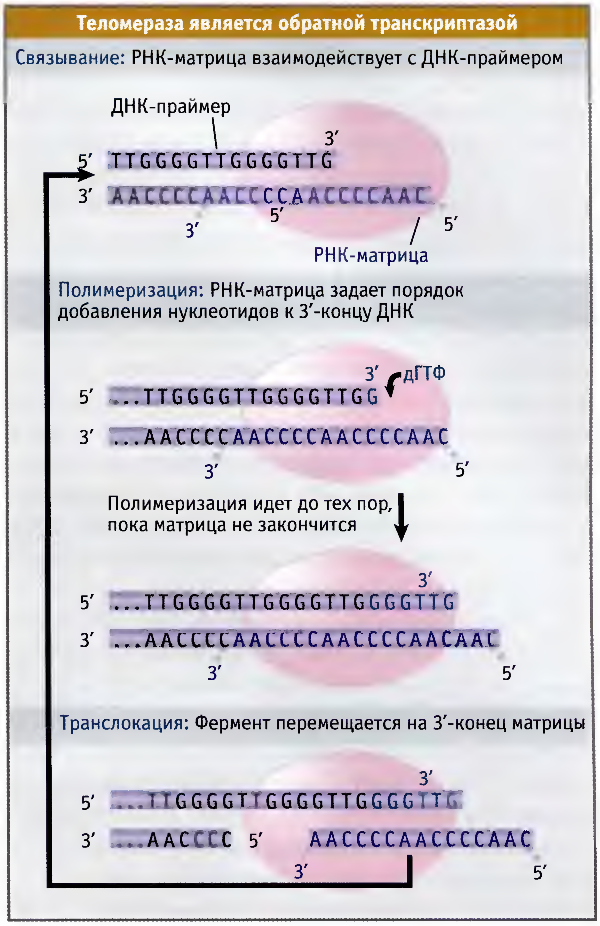


РИС. 28.31 Теломераза устраивается на субстрате за счет взаимодействия РНК-матрицы с затравкой – выступающей одиночной цепью ДНК. Теломераза добавляет к затравке нуклеотиды G и T в порядке, диктуемом структурой матрицы. После синтеза каждого нового повтора цикл работы фермента возобновляется

У теломеры есть две функции:

- Первая — это защищать конец хромосомы. Любой другой конец ДНК — например, возникший в результате двухцепочечного разрыва — это мишень (субстрат) для систем репарации. Теломера охраняет конец хромосомы от систем репарации; клетка отличает теломеру от других свободных концов.
- Вторая функция теломеры — позволять концу хромосомной ДНК удлиняться. Если он не удлиняется, то с каждым циклом репликации становится короче (потому что репликативный синтез ДНК не может начаться с самого начала матрицы).

Решение двух независимых функциональных задач обеспечивают белки, которые связываются с теломерой. У дрожжей эти задачи решаются разными наборами белков, но оба набора связываются с теломерой через один и тот же белок-адаптор Cdc13:

- Белок Stn1 защищает теломерную ДНК от деградации (если точнее, от излишнего распространения деградации С-А-богатой цепи, в норме необходимой, чтобы образовывался G-хвост).
- Фермент **теломераза** удлиняет G,T-богатую цепь. На активность теломеразы влияют два вспомогательных белка, контролирующих параметры наращивания (длину добавки и т.п.).

Теломераза использует 3′-ОН конец G-T-богатой цепи теломерной ДНК в качестве затравки для синтеза тандемных повторов TTGGGG; ей для этого нужны только дГТФ и дТТФ. Теломераза — это крупный рибонуклеопротеин, состоящий из РНК-праймера (кодируемого геном *TLC1*) и каталитически-активного белка (кодируемого геном *EST2*). РНК-компонент (который у *Tetrahymena* состоит из 159 оснований, а у *Euplotes* — из 192 оснований) содержит последовательность 15-22 оснований, идентичную двум повторам С-богатой цепи. Эта РНК служит матрицей для синтеза G-богатой повторяющейся последовательности теломерной ДНК. Белковый компонент теломеразы — это каталитическая субъединица, способная работать ДНК-полимеразой только если использует своего РНК-партнера (компонент теломеразы) в качестве матрицы.

Работа теломеразы, показанная на РИС. 28.31, состоит из дискретных шагов. РНК-матрица устанавливается на ДНК-затравку, каталитическая субъединица добавляет к затравке несколько нуклеотидов; затем, для начала следующего шага работы, фермент перемещается вдоль субстрата. Теломераза — это специализированная разновидность обратной транскриптазы — фермента, синтезирующего ДНК на матрице РНК (см. разд. 22.4 «ДНК ретровируса является продуктом обратной транскрипции»). Каким образом происходит синтез комплементарной (С-А-богатой)-цепи, не известно (можно домыслить, что 3′-конец G-T-богатой цепи может загнуться, как шпилька, и послужить затравкой).

Теломераза синтезирует добавочные повторы по отдельности, по одному на шаг, и сама не контролирует их число. Длина теломерной ДНК, выраженная в числе приращенных повторов, задается при участии других белков. Так, для дрожжей-мутантов по *EST1* или *EST3* характерны измененные длины теломер. Возможно, белки EST1 и EST3 связывают теломеразу и, контролируя доступ теломеразы к своему субстрату, тем самым влияют на длину теломерной ДНК. Белки,

связанные с теломерами в клетках млекопитающих, также идентифицированы, но об их функциях известно очень мало.

В хромосомах каждого вида организмов протяженность теломерной ДНК имеет свой характерный диапазон. У млекопитающих теломерные участки длинные (у людей, как правило, от 5 до 15 тпн), а у дрожжей — короткие (у *S. cerevisiae*, как правило, ~300 пн). Основа механизма, контролирующего их размеры, состоит в том, что, по мере укорачивания теломеры, вероятность стать субстратом теломеразы для ее ДНК возрастает. Как построен этот эффект (непрерывен ли, или дает себя знать только если длина теломеры падает ниже критического значения) не известно. Когда теломераза действует на теломеру, она может добавить к ее ДНК несколько повторов. Предоставленная самой себе, она добавит к субстрату один повтор, после чего диссоциирует; добавление нескольких повторяющихся единиц обусловлено другими белками, которые заставляют теломеразу предпринять более чем один раунд удлинения ДНК. На число добавляемых повторов не влияет длина самой теломеры; оно контролируется вспомогательными белками, ассоциированными с теломеразой.

Чтобы существовать, хромосоме необходим следующий минимум структур:

- Теломеры — для сохранности.
- Центромера — для поддержки сегрегации.
- Ориджин — для инициации репликации.

Соответствующие элементы ДНК используются как обязательные составляющие **искусственной дрожжевой хромосомы**. Это генно-инженерная конструкция, используемая как средство поддержания чужеродных последовательностей ДНК в дрожжевых клетках. Оказалось, что такие синтетические хромосомы стабильны только тогда, когда их длина превышает 20 (или даже 50) тпн.; природа этого ограничения непонятна. Все же умение конструировать синтетические хромосомы — это большой методический прорыв, позволяющий, в частности, исследовать природу сегрегации» в контролируемых условиях.

28.19 Теломеры принципиально важны для сохранности хромосом и выживания клеточных линий

В норме, пока клетки способны делиться, теломераза в них активна; после прохождения клеткой терминальной дифференцировки теломераза отключается. На РИС. 28.32 показано, что при мутации теломеразы у потомков клетки-мутанта с каждым новым делением теломеры становятся все короче и короче. Реальный пример последствий такой мутации у дрожжей приведен на



РИС. 28.32 Из-за генетического нарушения теломеразной активности теломеры хромосом от деления к делению постепенно укорачиваются. Потеря теломер приводит к разломам и хромосомным перестройкам

РИС. 28.33: теломерная ДНК укорачивается с 400 пн до нуля за ~120 поколений.

Утрата теломер весьма плачевна. Клетки, у которых длины теломер свелись к нулю, не способны успешно делиться: попытки делений обычно приводят к хромосомным перестройкам (разломам и транслокациям), а это сопряжено с повышенной частотой возникновения

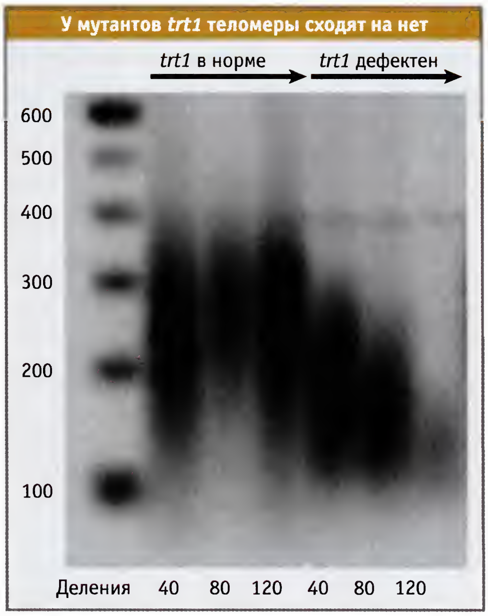


РИС. 28.33 В дрожжевых клетках дикого типа длина теломер поддерживается на уровне ~350 пн. У мутантов по гену *trt1*, кодирующему РНК-компонент теломеразы, теломеры быстро сокращаются до нуля. Воспроизведено по Nakamura, T. M., et al. 1997. Science. 277: 955–959. © 1997 AAAS. Фотография любезно предоставлена Thomas R. Cech and Toru Nakamura, University of Colorado

мутаций. У дрожжей это ассоциировано с потерей жизнеспособности и преобладанием в культуре стареющих клеток (имеющих удлиненную форму и постепенно вымирающих).

Некоторым клеткам удается перерасти стареющую культуру. Каким-то образом такие клетки поддерживают свои хромосомы в сохранности, пользуясь механизмом, альтернативным активности теломеразы. Выжившие клетки можно подразделить на две группы. В клетках первой группы хромосомы замкнулись в форме колец; теперь у них нет ни теломер, ни концов, и от теломеразы их сохранность не зависит. Выжившие клетки другой группы используют неравный кроссинговер, чтобы удлинять теломеры своих хромосом. Этот способ выживания проиллюстрирован на РИС. 28.34. Каждая теломера содержит очередь повторов ДНК, и когда гомологичные хромосомы становятся в пары, велика вероятность смещения теломерных участков их ДНК друг относительно друга. Рекомбинация между неверно состыкованными участками представляет собой неравный кроссинговер. Этот случай уже был показан на рис. 6.11: если длина одной рекомбинантной хромосомы увеличивается, то другая рекомбинантная хромосома при этом укорачивается.

В нормальных клетках неравный кроссинговер подавляется — из-за угрозы делеций, которую он создает. Подавлением кроссинговера между теломерами занимаются две клеточных системы. Одна из них — это белки, связывающиеся с теломерой. Так, у дрожжей частота событий рекомбинации между теломерами увеличивается при делеции гена *taz1*; этот ген кодирует белок, регулирующий теломеразную активность. Вторая система подавления неравного кроссинговера — это система репарации ошибочного спаривания. В дополнение к собственно коррекции неверно спаренных оснований, возникающих в ДНК, эта система подавляет рекомбинацию между плохо спаренными участками. Как показано на рис. 28.34, теломеры как раз относятся к таким участкам. Нарушения в этой системе репарации приводят к увеличению доли жизнеспособных клеток в дрожжевой культуре с дефицитом теломеразной активности. Выживание этих клеток обусловлено рекомбинацией между теломерными участками ДНК, удлиняющей теломеры некоторых хромосом.

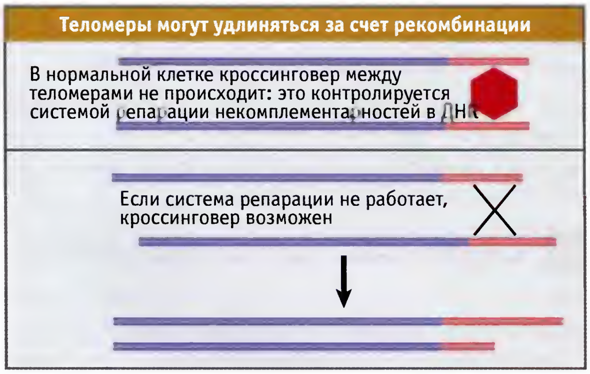


РИС. 28.34 В нормальных клетках обмен цепями (кроссинговер) теломерной ДНК подавлен благодаря работе определенных систем репарации. В клетках, где работа таких систем нарушена из-за мутаций, хромосомы могут обмениваться цепями теломерной ДНК. Событие неравного кроссинговера удлиняет теломеру одной из хромосом и позволяет этой хромосоме пережить отсутствие теломеразы

Эукариотические клетки, переведенные из состава ткани в культуру, через фиксированное число делений, как правило, претерпевают старение. Причина затухания их жизнеспособности — редукция теломер из-за спада экспрессии теломеразы. Клетки входят в кризис, от которого некоторые из них оправляются — обычно, ценой хромосомных перестроек, проистекших от незащищенности концов хромосом. Эти перестройки могут быть причиной возникновения мутаций, способствующих переходу клетки в опухолеродное состояние. Отсутствие экспрессии теломеразы у клеток в первичной культуре обусловлено инактивацией транскрипции ее гена; реактивация теломеразы — одна из стратегий выживания клеточных линий в условиях продолжительного культивирования. (Эта стратегия, конечно же, не являлась опцией для дрожжевых клеток в экспериментах, где ген теломеразы был делетирован.)

28.20 Резюме

Генетический материал всех организмов и вирусов принимает форму плотно упакованного нуклеопротеина. Геномы некоторых вирусов укладываются в подготовленный (предварительно собранный) капсид вириона; у других вирусов белковая оболочка собирается вокруг нуклеиновой кислоты. Бактериальный геном в сочетании с белками (20% по массе) образует нуклеоид; детали взаимодействия белков с ДНК не известны. ДНК бактериального нуклеоида организована в виде ~100 доменов, сверхспирализованность которых поддерживается независимо; плотность свободных сверхвитков в них составляет один на 100–200 пн. Хромосомы эукариот — равно как интерфазный хроматин — организованы в виде крупных петель. Каждая петля может быть независимым сверхспирализованным доменом. Основания петель соединяются с каркасами хромосом (в метафазе) или с ядерным матриксом (в интерфазе) через специфические сайты ДНК.

Транскрипционно-активные геномные последовательности входят в состав эухроматина — формации, к которой принадлежит большая часть хроматина интерфазного ядра. В составе другой формации — гетерохроматина — геномная ДНК упакована в 5–10 раз более компактно, чем в эухроматине; ДНК гетерохроматина

транскрипционно инертна. Во время клеточного деления весь хроматин переходит в плотно упакованное состояние, так что становятся различимы отдельные хромосомы. Существование воспроизводимой ультраструктуры хромосом доказывается возникновением в них характерных паттернов G-полос при обработке препаратов красителем Гимза. Полосы соответствуют очень крупным участкам генома (~10^7 пн) и могут быть использованы для картирования транслокаций или других хромосомных перестроек.

Хромосомы типа ламповых щеток у амфибий и политенные хромосомы у насекомых имеют нестандартную релаксированную структуру с коэффициентом упаковки <100. Набор политенных хромосом *D. melanogaster* подразделяется на ~5000 полос (не считая промежутков). Размер полос имеет разброс на порядок величины, средний размер составляет 25 тпн. Транскрипционно-активные участки могут быть визуализированы, как еще более релаксированные (вздутые) структуры — так называемые пуфы, материал которых как бы выдавлен за пределы осевой структуры хромосомы. Того же рода структурные изменения, но в меньшем масштабе, могут сопровождать транскрипцию в эухроматине.

Центромера (область хромосомы) несет кинетохор — структуру, ответственную за прикрепление хромосомы к веретену клеточного деления. Часто центромера находится в окружении гетерохроматина. Центромерные последовательности ДНК были четко идентифицированы только в дрожжах *S. cerevisiae*; там они состоят из коротких элементов. Консервативные элементы *CDE-I* и *CDE-III* связываются белковыми комплексами CBF1 и CBF3, соответственно; более длинный A-T-богатый элемент *CDE-II*, расположенный между *CDE-I* и *CDE-III*, связывается белком Cse4. Этими взаимодействиями обуславливается возникновение центромеры, как специализированной структуры в составе хроматина, с которой связывается другая группа белков, опосредующая связь с микротрубочками веретена деления.

Теломеры придают стабильность концам хромосом. Почти во всех известных случаях теломерная ДНК состоит из многократных повторов, где одна цепь имеет вид $C_n(A/T)_m$; $n > 1$; $m = 1,2,3$ или 4. Комплементарная цепь $G_n(T/A)_m$ имеет выступающий одноцепочечный конец; он служит затравкой для наращивания теломерной ДНК. Наращивание теломерной ДНК осуществляет фермент теломераза — это рибонуклеопротеин, чей РНК-компонент служит матрицей для синтеза G-богатой цепи теломерной ДНК. Наращивание позволяет клетке справиться с проблемой репликации краев линейного дуплекса ДНК. Теломера стабилизирует конец хромосомы еще и тем, что аннулирует его как субстрат для репарации: одноцепочечный концевой выступ теломерной ДНК $G_n(T/A)_m$ производит замещение гомологичного участка в вышележащем дуплексе, и свободные концы хромосомной ДНК оказываются спрятанными в составе образовавшейся петли.

Литература

28.2 Вирусный геном упакован в оболочку

Обзор

Black, L. W. (1989). DNA packaging in dsDNA bacteriophages. *Annu. Rev. Immunol.* 43, 267–292.

Butler, P. J. (1999). Self-assembly of tobacco mosaic virus: the role of an intermediate aggregate in generating both specificity and speed. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 354, 537–550.

Klug, A. (1999). The tobacco mosaic virus particle: structure and assembly. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 354, 531–535.

Mindich, L. (2000). Precise packaging of the three genomic segments of the double-stranded-РНК bacteriophage phi6. *Microbiol. Mol. Biol. Rev.* 63, 149–160.

Исследование

Caspar, D. L. D. and Klug, A. (1962). Physical principles in the construction of regular viruses. *Cold Spring Harbor Symp. Quant. Biol.* 27, 1–24.

de Beer, T., Fang, J., Ortega, M., Yang, Q., Maes, L., Duffy, C., Berton, N., Sippy, J., Overduin, M., Feiss, M., and Catalano, C. E. (2002). Insights into specific DNA recognition during the assembly of a viral genome packaging machine. *Mol. Cell* 9, 981–991.

Dube, P., Tavares, P., Lurz, R., and van Heel, M. (1993). The portal protein of bacteriophage SPP1: a DNA pump with 13-fold symmetry. *EMBO J.* 12, 1303–1309.

Fraenkel-Conrat, H. and Williams, R. C. (1955). Reconstitution of active tobacco mosaic virus from its inactive protein and nucleic acid components. *Proc. Natl. Acad. Sci. USA* 41, 690–698.

Jiang, Y. J., Aerne, B. L., Smithers, L., Haddon, C., Ish-Horowicz, D., and Lewis, J. (2000). Notch signalling and the synchronization of the somite segmentation clock. *Nature* 408, 475–479.

Zimmern, D. (1977). The nucleotide sequence at the origin for assembly on tobacco mosaic virus РНК. *Cell* 11, 463–482.

Zimmern, D. and Butler, P. J. (1977). The isolation of tobacco mosaic virus РНК fragments containing the origin for viral assembly. *Cell* 11, 455–462.

28.3 Бактериальный геном образует нуклеоид

Обзор

Brock, T. D. (1988). The bacterial nucleus: a history. *Microbiol. Rev.* 52, 397–411.

Drlica, K. and Rouviere-Yaniv, J. (1987). Histone-like proteins of bacteria. *Microbiol. Rev.* 51, 301–319.

28.4 Бактериальный геном сверхспирализован

Обзор

Hatfield, G. W. and Benham, C. J. (2002). DNA topology-mediated control of global gene expression in *Escherichia coli*. *Annu. Rev. Genet.* 36, 175–203.

Исследование

Pettijohn, D. E. and Pfenninger, O. (1980). Supercoils in prokaryotic DNA restrained *in vitro*. *Proc. Natl. Acad. Sci. USA* 77, 1331–1335.

Postow, L., Hardy, C. D., Arsuaga, J., and Cozzarelli, N. R. (2004). Topological domain structure of the *Escherichia coli* chromosome. *Genes Dev.* 18, 1766–1779.

28.8 У хромосом проявляются полосатые паттерны окрашивания

Исследование

IntePHKtional Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860–921.

Saccone, S., De Sario, A., Wiegant, J., Raap, A. K., Della Valle, G., and BePHKrdi, G. (1993). Correlations between isochores and chromosomal bands in the human genome. *Proc. Natl. Acad. Sci. USA* 90, 11929–11933.

Venter, J. C. et al. (2001). The sequence of the human genome. *Science* 291, 1304–1350.

28.12 Эукариотические хромосомы служат клетке приспособлениями для сегрегации генетического материала

Обзор

Hyman, A. A. and Sorger, P. K. (1995). Structure and function of kinetochores in budding yeast. *Annu. Rev. Cell Dev. Biol.* 11, 471–495.

28.13 Центромеры содержат повторяющиеся последовательности ДНК

Обзор

Wiens, G. R. and Sorger, P. K. (1998). Centromeric chromatin and epigenetic effects in kinetochore assembly. *Cell* 93, 313–316.

Исследование

Copenhaver, G. P. et al. (1999). Genetic definition and sequence analysis of *Arabidopsis* centromeres. *Science* 286, 2468–2474.

Haaf, T., Warburton, P. E., and Willard, H. F. (1992). Integration of human alpha-satellite DNA into simian chromosomes: centromere protein binding and disruption of normal chromosome segregation. *Cell* 70, 681–696.

Sun, X., Wahlstrom, J., and Karpen, G. (1997). Molecular structure of a functional *Drosophila* centromere. *Cell* 91, 1007–1019.

28.14 Центромеры *S. cerevisiae* несут короткие последовательности ДНК

Обзор

Blackburn, E. H. and Szostak, J. W. (1984). The molecular structure of centromeres and telomeres. *Annu. Rev. Biochem.* 53, 163–194.

Clarke, L. and Carbon, J. (1985). The structure and function of yeast centromeres. *Annu. Rev. Genet.* 19, 29–56.

Исследование

Fitzgerald-Hayes, M., Clarke, L., and Carbon, J. (1982). Nucleotide sequence comparisons and functional analysis of yeast centromere DNAs. *Cell* 29, 235–244.

28.15 Центромера связывается с белковым комплексом

Обзор

Kitagawa, K. and Hieter, P. (2001). Evolutionary conservation between budding yeast and human kinetochores. *Nat. Rev. Mol. Cell Biol.* 2, 678–687.

Исследование

Lechner, J. and Carbon, J. (1991). A 240 kd multisubunit protein complex, CBF3, is a major component of the budding yeast centromere. *Cell* 64, 717–725.

Meluh, P. B. and Koshland, D. (1997). Budding yeast centromere composition and assembly as revealed by *in vitro* cross-linking. *Genes Dev.* 11, 3401–3412.

Meluh, P. B. et al. (1998). Cse4p is a component of the core centromere of *S. cerevisiae*. *Cell* 94, 607–613.

Ortiz, J., Stemmann, O., Rank, S., and Lechner, J. (1999). A putative protein complex consisting of Ctf19, Mcm21, and Okp1 represents a missing link in the budding yeast kinetochore. *Genes Dev.* 13, 1140–1155.

28.16 Теломеры содержат простые повторяющиеся последовательности ДНК

Обзор

Blackburn, E. H. and Szostak, J. W. (1984). The molecular structure of centromeres and telomeres. *Annu. Rev. Biochem.* 53, 163–194.

Zakian, V. A. (1989). Structure and function of telomeres. *Annu. Rev. Genet.* 23, 579–604.

Исследование

Wellinger, R. J., Ethier, K., Labrecque, P., and Zakian, V. A. (1996). Evidence for a new step in telomere maintenance. *Cell* 85, 423–433.

28.17 Теломеры опечатывают концы хромосом

Исследование

Griffith, J. D. et al. (1999). Mammalian telomeres end in a large duplex loop. *Cell* 97, 503–514.

Henderson, E., Hardin, C. H., Walk, S. K., Tinoco, I., and Blackburn, E. H. (1987). Telomeric oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs. *Cell* 51, 899–908.

Karlseder, J., Broccoli, D., Dai, Y., Hardy, S., and de Lange, T. (1999). p53- and ATM-dependent apoptosis induced by telomeres lacking TRF2. *Science* 283, 1321–1325.

Parkinson, G. N., Lee, M. P., and Neidle, S. (2002). Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* 417, 876–880.

van Steensel, B., Smogorzewska, A., and de Lange, T. (1998). TRF2 protects human telomeres from end-to-end fusions. *Cell* 92, 401–413.

Williamson, J. R., Raghuraman, K. R., and Cech, T. R. (1989). Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell* 59, 871–880.

28.18 Теломеры синтезируются рибонуклеопротеиновым ферментом

Обзор

Blackburn, E. H. (1991). Structure and function of telomeres. *Nature* 350, 569–573.

Blackburn, E. H. (1992). Telomerases. *Annu. Rev. Biochem.* 61, 113–129.

Collins, K. (1999). Ciliate telomerase biochemistry. *Annu. Rev. Biochem.* 68, 187–218.

Smogorzewska, A. and de Lange, T. (2004). Regulation of telomerase by telomeric proteins. *Annu. Rev. Biochem.* 73, 177–208.

Zakian, V. A. (1995). Telomeres: beginning to understand the end. *Science* 270, 1601–1607.

Zakian, V. A. (1996). Structure, function, and replication of *S. cerevisiae* telomeres. *Annu. Rev. Genet.* 30, 141–172.

Исследование

Greider, C. and Blackburn, E. H. (1987). The telomere terminal transferase of *Tetrahymena* is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell* 51, 887–898.

Murray, A., and Szostak, J. W. (1983). Construction of artificial chromosomes in yeast. *Nature* 305, 189–193.

Pennock, E., Buckley, K., and Lundblad, V. (2001). Cdc13 delivers separate complexes to the telomere for end protection and replication. *Cell* 104, 387–396.

Shippen-Lentz, D. and Blackburn, E. H. (1990). Functional evidence for an PHK template in telomerase. *Science* 247, 546–552.

Teixeira, M. T., Arneric, M., Sperisen, P., and Lingner, J. (2004). Telomere length homeostasis is achieved via a switch between telomerase-extendible and -nonextendible states. *Cell* 117, 323–335.

28.19 Теломеры принципиально важны для сохранности хромосом и выживания клеточных линий

Исследование

Hackett, J. A., Feldser, D. M., and Greider, C. W. (2001). Telomere dysfunction increases mutation rate and genomic instability. *Cell* 106, 275–286.

Nakamura, T. M., Cooper, J. P., and Cech, T. R. (1998). Two modes of survival of fission yeast without telomerase. *Science* 282, 493–496.

Nakamura, T. M., Morin, G. B., Chapman, K. B., Weinrich, S. L., Andrews, W. H., Lingner, J., Harley, C. B., and Cech, T. R. (1997). Telomerase catalytic subunit homologs from fission yeast and human. *Science* 277, 955–959.

Rizki, A. and Lundblad, V. (2001). Defects in mismatch repair promote telomerase-independent proliferation. *Nature* 411, 713–716.

Нуклеосомы

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

29.1 Введение

Хроматин имеет компактную организацию, внутри которой большинство последовательностей ДНК из-за своей структурированности недосягаемы и функционально неактивны. Среди этой массы попадаются активные последовательности, но их меньшинство. Как устроен хроматин и в чем заключается различие между неактивными и активными последовательностями — с точки зрения его структуры? Чрезвычайно высокий коэффициент упаковки ДНК в хроматине наводит на мысль о наличии в нем структурной *иерархии*.

Фундаментальная структурная единица хроматина у всех эукариот одинакова. **Нуклеосома** содержит ~200 пн ДНК; этот отрезок двойной спирали связан с октамером небольших основных белков, образуя структуру, похожую на бусину. Эти белки, **гистоны**, образуют сердцевину бусины, по поверхности которой проходит ДНК. Нуклеосомы — инвариантный компонент эухроматина и гетерохроматина интерфазных ядер, а также митотических хромосом. Нуклеосомы — это первый уровень организации хроматина, с коэффициентом упаковки геномной ДНК ~6. Состав и структура нуклеосом подробно изучены.

Второй уровень организации хроматина — это укладка серии нуклеосом в виде спирального массива с образованием волокна (фибриллы) диаметром ~30 нм; такое волокно есть и в интерфазном хроматине, и в митотических хромосомах (см. рис. 28.11). При этом коэффициент упаковки ДНК доводится до ~40 (в интерфазном хроматине). Структура таких волокон изучена не очень хорошо; известно, что ее поддержание требует участия дополнительных белков.

Окончательный коэффициент упаковки геномной ДНК определяется третьим уровнем организации хроматина, в рамках которого происходит упаковка самого 30-нанометрового волокна. Минимальный общий коэффициент упаковки, ~1000, характерен для эухроматина; исходя из этого минимума, значение коэффициента упаковки циклически колеблется, достигая своего максимума, ~10 000, в митотических хромосомах. В составе гетерохроматина, в интерфазе и при митозе, ДНК упакована с постоянным коэффициентом ~10 000.

Умозрительно, изменение структуры хроматина предполагает или ассоциацию с ним дополнительных белков, или модификацию конститутивных хромосомных белков. Конкретные мишени и рычаги контроля циклических изменений структуры хроматина доподлинно не известны. Репликация и транскрипция требуют расплетания ДНК и поэтому должны сопровождаться роспуском структуры хроматина — чтобы соответствующие белки получили доступ к ДНК и возможность ма-

нипулировать ею; судя по всему, изменения затрагивают организацию хроматина на всех трех ее уровнях.

Когда в хроматине идет репликация, нуклеосомы должны быть воспроизведены для обоих дочерних дуплексов геномной ДНК. Задавшись вопросом сборки нуклеосомы, следует поинтересоваться и тем, что случается при репликации ДНК с другими белками, присутствующими в хроматине. Репликация возмущает структуру хроматина. При этом, с одной стороны, возникает проблема поддержания формы участков со специфической структурой; с другой стороны, возникает возможность для изменения структуры.

Массовая доля белков в хроматине чуть ли не в два раза больше массовой доли ДНК. Примерно половину массы хроматиновых белков следует отнести на счет нуклеосом. Масса РНК составляет менее 10% от массы ДНК; значительная часть РНК — это новосинтезированные транскрипты, все еще сохраняющие связь с матрицей ДНК.

К **негистоновым** белкам относятся все белки хроматина за вычетом гистонов. По сравнению с гистонами, негистоновые белки более вариабельны между разными типами тканей и видами организмов; их массовая доля в хроматине чуть меньше, чем доля гистонов. Негистоновые белки гораздо разнообразнее гистонов, так что в хроматине каждый отдельно взятый негистоновый белок присутствует в количестве, гораздо меньшем, нежели любой из гистонов.

В обязанности негистоновых белков входит контроль экспрессии генов и структуры хроматина высшего порядка. Так, РНК-полимеразу можно рассматривать как негистоновый белок. Особая группа негистоновых белков — белки высокой подвижности (high mobility group proteins, HMG proteins) — включает в себя ряд факторов транскрипции. В работе с другими негистоновыми белками остро стоит проблема контаминации их препаратов другими ядерными белками. Методическая задача очистки негистоновых белков, отвечающих в хроматине за структуры высших порядков, до сих пор в совершенстве не решена.



РИС. 29.1 Клеточные ядра, у которых повреждена оболочка, выпускают из себя хроматин: он состоит из компактно организованных частиц. Мерка равняется 100 нм. Воспроизведено по *Cell*, vol. 4, Oudet, P., et al., Electron microscopic . . . , pp. 281–300. Авторские права 1975, с разрешения Elsevier. Фотография любезно предоставлена Pierre Chambon

29.2 Любой хроматин состоит из субъединиц — нуклеосом

Ключевые понятия:

- Микрококковая нуклеаза разрезает хроматин на отдельные нуклеосомы — частицы с коэффициентом седиментации 11S.
- Нуклеосома содержит ~200 пн ДНК и по две молекулы каждого из гистонов сердцевины (H2A, H2B, H3 и H4).
- ДНК обернута вокруг наружной поверхности гистонового октамера.

Если интерфазные ядра суспендировать в растворе низкой ионной силы, то они раздуваются и лопаются, выпуская из себя волокна (фибриллы) хроматина. На РИС. 29.1 показано ядро, подвергнутое такому лизису, с устремленными из него наружу фибриллами хроматина. Некоторые участки хроматина, как можно заметить, сохранили плотность упаковки; участки, которые оказались растянутыми, со всей очевидностью, состоят из дискретных частиц — это и есть нуклеосомы. На особенно



РИС. 29.2 Отдельные нуклеосомы можно получить, обработав хроматин микрококковой нуклеазой. Мерка равняется 100 нм. Воспроизведено по *Cell*, vol. 4, Oudet, P., et al., Electron microscopic . . . , pp. 281–300. Copyright 1975, с разрешения Elsevier. Фотография любезно предоставлена Pierre Chambon

растянутых участках хроматина видно, что нуклеосомы соединены тонкой нитью; она представляет собой двухцепочечную ДНК. *Свободный дуплекс ДНК непрерывной нитью пробегает серию частиц.*

Отдельные нуклеосомы можно получить, обработав хроматин эндонуклеазой (**микрококковой нуклеазой**), разрезающей ДНК в местах соединения нуклеосом между собой. В начале такой обработки из хроматина высвобождаются группы частиц, которые затем разделяются на одиночные нуклеосомы. Их можно увидеть на РИС. 29.2 как компактные частицы. Седиментируют они с коэффициентом ~11S.

Нуклеосома содержит около 200 пн ДНК, ассоциированной с гистоновым октамером, который включает в себя по две молекулы гистонов H2A, H2B, H3 и H4. Эти белки называют **гистонами сердцевины**. Их ассоциация схематично изображена на РИС. 29.3. Эта модель хорошо объясняет стехиометрию гистонов сердцевины в составе хроматина, где H2A, H2B, H3 и H4 присутствуют в эквимолярных количествах, причем по две молекулы каждого из них приходится на 200 пн ДНК.

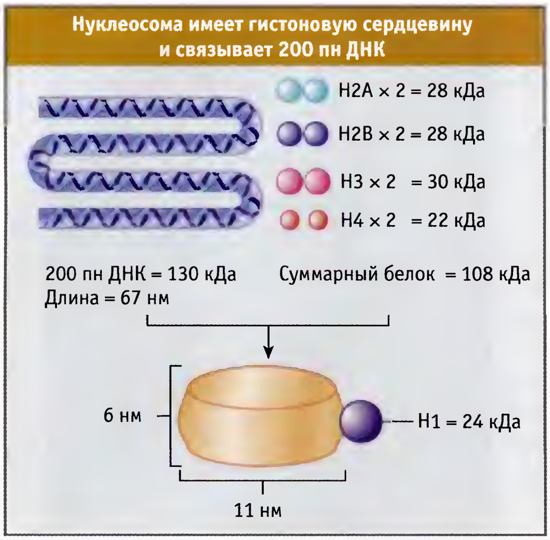


РИС. 29.3 Нуклеосома состоит из приблизительно равных масс ДНК и гистонов (с учетом H1). Расчетная масса нуклеосомы составляет 262 кДа



РИС. 29.4 Можно представить нуклеосому как цилиндр, по поверхности которого проходят два сверхвитка ДНК

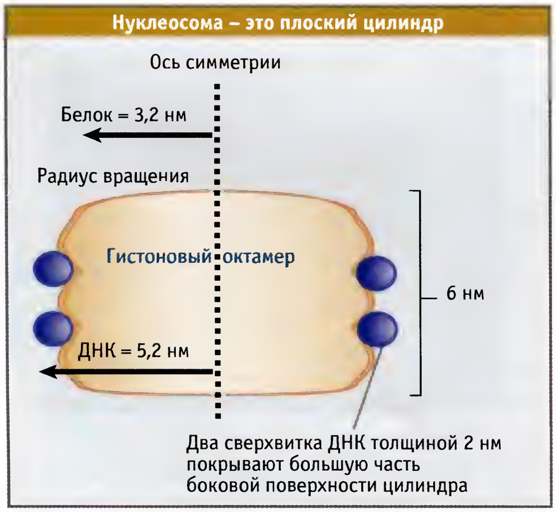


РИС. 29.5 В нуклеосоме два сверхвитка ДНК уложены близко друг к другу

Гистоны H3 и H4 известны как одни из самых консервативных белков. Это обстоятельство указывает на идентичность их функций у всех эукариот. Гистоны H2A и H2B также обнаруживаются у всех эукариот, но с существенными видоспецифичными вариациями аминокислотных последовательностей.

Гистон H1 — это собирательное имя близкородственных белков, отличающихся у разных видов организмов и в зависимости от типа ткани. У H1 роль другая, чем у гистонов сердцевины. Из хроматина, где H1 присутствует в количестве, половинном от количества любого из гистонов сердцевины, он экстрагируется легче, чем гистоны средцевины (обычно это делается разбавленным [0,5 М] раствором соли). К тому же, *H1 можно удалить из нуклеосомы, не нарушив ее структуры, указывая на то, что он расположен снаружи частицы.*

По своей форме нуклеосома соответствует диску, или уплощенному цилиндру, диаметром 11 нм и высотой 6 нм. Длина ДНК в составе частицы, очень приблизительно, соответствует двум окружностям (~34 нм) октамера гистонов сердцевины. ДНК ложится на октамер симметрично. На РИС. 29.4 ее траектория схематично показана как спиральная обмотка из двух витков вокруг цилиндрического октамера сердцевины нуклеосомы. Заметим, что ДНК «входит» в нуклеосому (точнее, в центральную часть нуклеосомы, состоящую из гистоновой сердцевины, дважды обернутой ДНК) и «покидает» ее в точках, расположенных близко друг к другу. Гистон H1, возможно, находится рядом с этими точками (см. разд. 29.4 «Все нуклеосомы устроены одинаково»).

Если рассмотреть эту модель в виде поперечного среза через нуклеосому, как показано на РИС. 29.5, то можно увидеть, что два витка ДНК на поверхности гистонового октамера лежат близко друг к другу. Из 6 нм высоты цилиндра 4 нм занято двумя витками ДНК (с диаметром двойной спирали, как известно, 2 нм).



РИС. 29.6 Из-за тесного расположения сверхвитков нуклеосомной ДНК ее сайты перегруппированы друг относительно друга

Такая конфигурация витков ДНК в составе нуклеосомы имеет возможное функциональное следствие. Один оборот ДНК вокруг сердцевины нуклеосомы состоит из ~80 пн, и две точки, разделенные 80 пн двойной спирали, в нуклеосоме могут оказаться сближенными (как показано на РИС. 29.6).

29.3 Геномная ДНК свернута в виде массивов нуклеосом

Ключевые понятия:

- После обработки хроматина микрококковой эндонуклеазой >95% ДНК можно собрать с нуклеосом или их мультимеров.
- Длина ДНК, приходящейся на одну нуклеосому в клетках разного происхождения, варьирует в диапазоне от 154 до 260 пн.

Если хроматин обработать ферментом микрококковой нуклеазой, то геномная ДНК будет нарезана в виде фрагментов, соответствующих некоторой фиксированной единице длины, умноженной на целое число. Фракционирование такого образца ДНК гель-электрофорезом дает «лестницу» из фрагментов разной длины, представленную на РИС. 29.7. Обычно такие «лестницы» имеют около 10 «ступенек», а единица длины, определяемая как расстояние между соседними «ступеньками», равняется ~200 пн.

Фракционирование нуклеосом в градиенте сахарозы дает серию дискретных пиков, которые соответствуют их мономерам, димерам, тримерам и т.д. Если из разделенных фракций экстрагировать ДНК и, не смешивая, подвергнуть электрофорезу, то каждой фракции будет соответствовать одна полоса уникальной длины, эквивалентная одной из ступенек микрококково-нуклеазной «лестницы». Это соответствие между «ступеньками лестницы» на электрофореграмме и седиментировавшими группами из нескольких связанных нуклеосом показано на РИС. 29.8. Мономерная нуклеосома содержит одну единицу длины ДНК, димер нуклеосом — две единицы, и т.д.

В каждой «ступеньке лестницы» представлена ДНК, связанная с конкретным целым числом нуклеосом. *Выявление в каком-либо хроматине «лестницы» с шагом 200 пн можно расценивать как индикатор организации его ДНК в виде нуклеосом.* «Лестница» возникает уже тогда, когда в форму, растворимую в кислой среде (то есть деградированную нуклеазой до небольших фрагментов) перейдет всего 2% ядерной ДНК. Значит, *небольшая доля ядерной ДНК атакуется нуклеазой более активно; должно быть, ею представлены особо уязвимые участки генома.*

Когда хроматин выходит из ядра, подвергнутого лизису, можно разглядеть серии нуклеосом, соединенных свободной ДНК, как бусы — ниткой. С другой стороны, необходимость в плотной упаковке ДНК *in vivo* говорит о том, что, скорее всего, свободной ДНК в ядре не так уж много (или вовсе нет).

Это соображение подтверждается тем фактом, что, после деградации микрококковой нуклеазой, *>95% ДНК хроматина можно собрать* и «уложить» электрофорезом *в виде «лестницы» с шагом 200 пн.* Следовательно, почти вся ДНК ядра там присутствует в форме нуклеосом. Нуклеосомы в ядрах *in vivo*, судя по всему, расположены тесно, так что ДНК прямо переходит из одной из них на другую. Свободная ДНК появляется, вероятно, лишь в препаратах, в связи с потерями гистоновых октамеров при ее выделении.

Длина ДНК, приходящаяся на одну нуклеосому, может отклоняться от типового значения в 200 пн. В клетках конкретного типа соблюдается характерное среднее значение (±5 пн). Чаще всего оно попадает между 180 и 200 пн, хотя известны особые случаи: например, 154 пн у

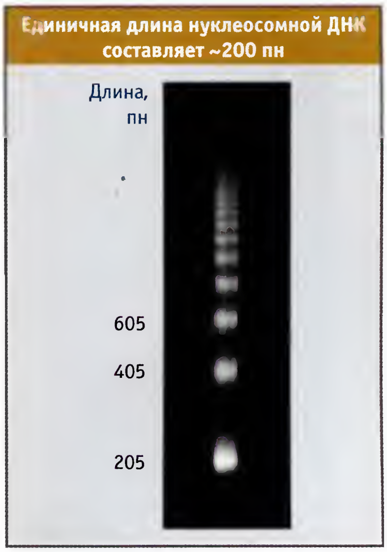


РИС. 29.7 Микрококковая нуклеаза расщепляет хроматин на мультимеры нуклеосом. ДНК этих мультимеров, разделенная гель-электрофорезом по длине фрагментов, предстает в виде «лестницы» из полос. Фотография любезно предоставлена Markus Noll, Universität Zürich

некоторых грибов или 260 пн в сперматозоидах морского ежа. Для разных тканей взрослого организма средние значения могут отличаться; отличия также могут быть для разных частей генома внутри одного клеточного типа. Отклонения от среднегеномного значения длины отрезка ДНК в отдельно взятой нуклеосоме свойственны тандемно повторяющимся последовательностям — таким как кластеры 5S РНК.

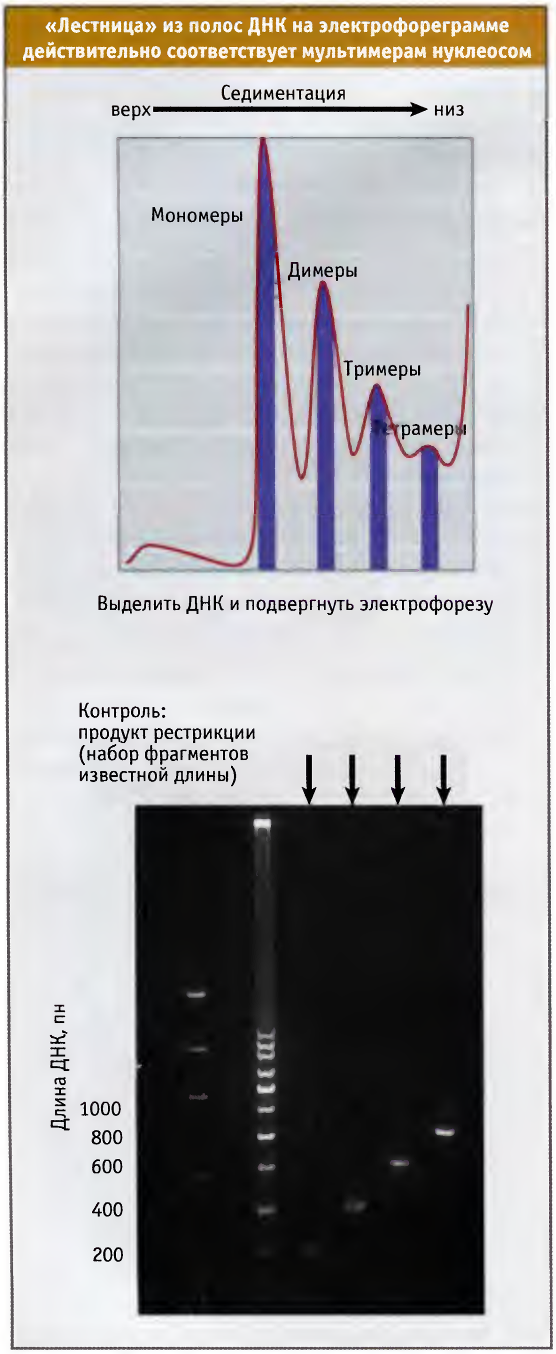


Рис. 29.8 Каждый мультимер нуклеосом содержит соответствующее число единичных длин нуклеосомной ДНК. Фотография любезно предоставлена Jan Kieleczawa, Wyeth Research

29.4 Все нуклеосомы устроены одинаково

Ключевые понятия:

- **По чувствительности к нуклеазе ДНК каждой нуклеосомы можно подразделить на две области — центральную и связующую.**
- **ДНК центральной области, протяженностью 146 пн, может быть снята с центральных частиц, образующихся при длительной обработке хроматина микрококковой нуклеазой.**
- **Связующая ДНК, протяженностью от 8 до 114 пн, уязвима для нуклеазы с самого начала обработки.**
- **Вариации суммарной протяженности ДНК одной нуклеосомы определяются колебаниями размеров связующей области.**
- **Гистон H1 ассоциирован со связующей ДНК и может находиться поблизости от точек «входа» и «выхода» (границ центральной области ДНК, ассоциированной с гистоновым октамером).**

Вариабельность длины ДНК, приходящейся на одну нуклеосому, в зависимости от источника ее происхождения, накладывается, тем не менее, на общность строения всех нуклеосом. При ассоциации ДНК с гистоновым октамером образуется центральная частица, содержащая 146 пн ДНК — вне зависимости от суммарной длины ДНК, приходящейся на данную нуклеосому. Изменчивость суммарной длины ДНК в нуклеосоме проецируется на постоянство структуры ее центральной части.

Получить центральную частицу можно, воздействуя микрококковой нуклеазой на нуклеосому-мономер. Если «напустить» фермент на интактный хроматин, то он разрежет ДНК между нуклеосомами. Если ферменту, после того как он нарезал хроматин на нуклеосомы-мономеры, позволить действовать дальше, то он расщепит часть ДНК из состава обособившихся нуклеосом. Эта реакция будет представлять собой «сострtriгание» краев нуклеосомной ДНК.

Уменьшение длины нуклеосомной ДНК происходит в форме дискретных шагов, это следует из анализа электрофореграммы на рис. 29.9. Исходная длина ДНК в нуклеосомах мономерах из ядер клеток печени крысы составляет 205 пар оснований. Если вести реакцию дальше, то через некоторое время можно обнаружить фрагменты ДНК, длина которых сократилась до ~165 пар оснований. Дальнейшее сокращение доводит длину фрагментов до размеров центральной области (входящей в состав центральной частицы), длина которой составляет 146 пар оснований. Эта центральная область довольно-таки стабильна, хотя если реакцию, в более жестких условиях, довести «до упора», то образуется так называемый «продукт предельного расщепления». В этом продукте самые длинные фрагменты — те же 146 пн центральной области, в то время как другие фрагменты могут быть гораздо короче — вплоть до 20 пн.

Такой анализ доказывает правомочность разделения нуклеосомной ДНК на две области:

- центральную область, с инвариантной длиной 146 пн и относительно устойчивую к расщеплению нуклеазами, и
- связующую область, пограничную с другими нуклеосомами. Ее длина варьирует от 8 до 114 пн.

Единичный отрезок нуклеосомной ДНК имеет центральную часть

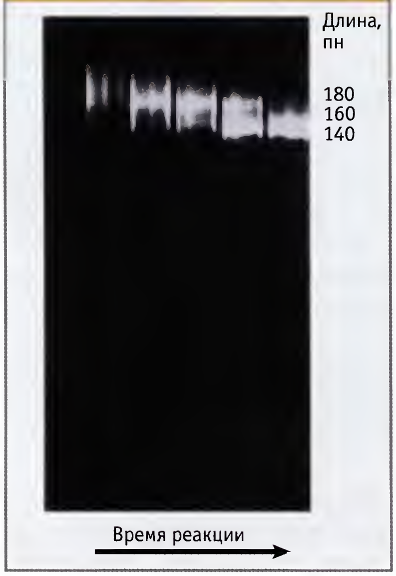


РИС. 29.9 Длина мономеров нуклеосомной ДНК поэтапно уменьшается до 146 пн. Фотография любезно предоставлена Roger Kornberg, Stanford University School of Medicine

Четкость полосы, которая возникает на электрофореграмме после самого первого этапа расщепления ДНК хроматина, говорит о том, что действию фермента исходно доступен лишь очень ограниченный участок — небольшой сегмент связующей области. (Если бы уязвимой была вся связующая область, то полоса выглядела бы размытой, потому что состояла бы из фрагментов с разбросом по длине от 146 до >200 пн.) Но как только разрез внесен, вся связующая область ДНК нуклеосомы-мономера становится уязвимой и может быть довольно быстро удалена последующим действием фермента. Связь между этапами расщепления нуклеосомной ДНК изображена на РИС. 29.10.

Свойства центральных частиц подобны свойствам самих нуклеосом, хотя по своим размерам центральные частицы несколько меньше. Все же по форме и размерам они похожи на нуклеосомы; это показывает, что геометрия частицы задается взаимодействием между ДНК и белковым октамером центральной частицы. Технически, центральные частицы как гомогенную популяцию (однородных) объектов выделить легче. Поэтому в исследованиях на предмет структуры часто предпочитают использовать именно их, а не препараты цельных нуклеосом. (Дело в том, что избежать «состригания» краев нуклеосомной ДНК при выделении очень сложно.)

В чем заключается физическая сущность различия между центральной и связующей областями нуклеосомной ДНК? Да, в общем-то, ни в чем. *Эти термины (названия областей) были введены как рабочие определения, в целях описания участков нуклеосомной ДНК с точки зрения их относительной чувствительности к обработке нуклеазой,* безотносительно их реальной структуры. Все же, как оказалось, большая часть центральной области имеет высокую плотность искривлений, в то время как края центральной области и вся связующая область — более расправленные (см. разд. 29.5 «На поверхности нуклеосомы структура ДНЕ неоднородна»).

Оформление связующей ДНК как области всецело опирается на факторы, внешние по отношению к гистонам сердцевины. Эксперименты по воссозданию (самосборке) нуклеосом *in vitro* показали, что гистонам присуща способность к организации ДНК в центральные частицы, но из таких частиц не образуются нуклеосомы, которые разбивали бы ДНК на отрезки правильной длины. По-видимому, важным фактором такого разбиения *in vivo* является сверхспирализованность ДНК; на протяженность связующих областей ДНК в сериях нуклеосом влияют гистон H1 и/или негистоновые белки. В сборке нуклеосом из гистонов и ДНК *in vivo* участвуют также «белки сборки», которые сами не входят в состав нуклеосомы (см. разд. 29.9 «Воспроизводство хроматина требует сборки нуклеосом»).

Где, по отношению к центральной частице, располагается молекула гистона H1? Во время деградации нуклеосом-мономеров (путем расщепления их ДНК) H1 теряется, причем потеря происходит на этапе деградации промежуточного продукта (в котором 165 пн ДНК) до центральной частицы с 146 пн ДНК. Это говорит о том, что молекула H1 может быть ассоциирована с тем сегментом ДНК-связующей области, который непосредственно примыкает к центральной области.

Если H1 ассоциирован со связующей областью, он может «опечатать» нуклеосомную ДНК, связываясь с участками, в которых нуклеинованя кислота входит и покидает нуклеосому (см. рис. 29.4). Представление о расположении молекул H1 на стыках смежных нуклеосом хорошо согласуется с известным лабораторным опытом: удалить H1 из хроматина проще, чем другие гистоны, а хроматин, обедненный гистоном H1, «растворяется» лучше. Кроме того, если хроматиновую фибриллу освободить от H1, то растянуть ее как «бусы на нитке» гораздо легче.

Микрококковая нуклеаза «подстригает» нуклеосомы до центральных частиц

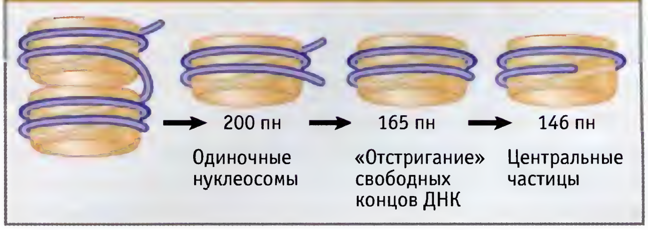


РИС. 29.10 Микрококковая нуклеаза режет ДНК между нуклеосомами. Одиночные нуклеосомы содержат по ~200 пн ДНК. «Отстригание» свободных концов ДНК при дальнейшей инкубации одиночных нуклеосом с той же микрококковой нуклеазой дает уменьшение длины мономеров нуклеосомной ДНК: сперва до ~165 пн, и, наконец, до ~146 пн.

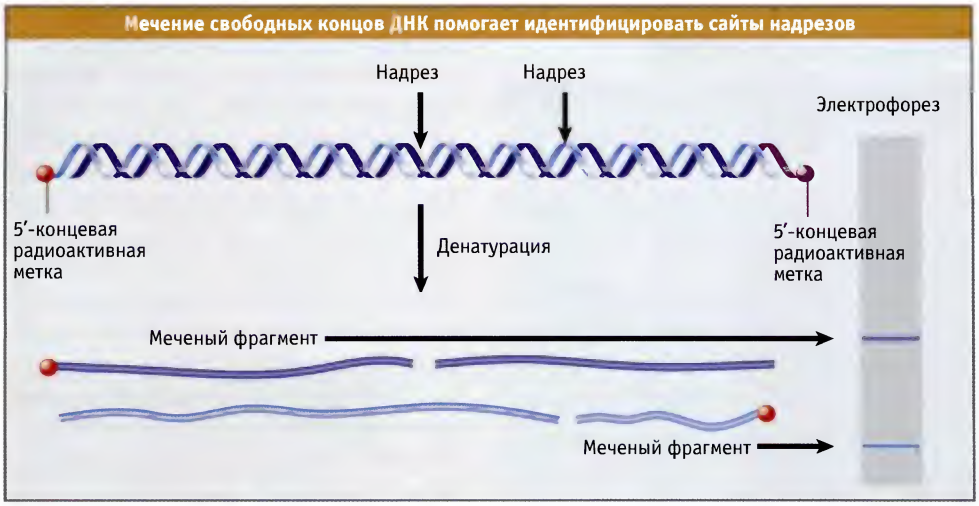


РИС. 29.11 Положение надреза в дуплексе можно идентифицировать, пометив свободный конец ДНК (например 5′-конец). Перед авторадиографией ДНК денатурируют. Тогда размер меченого фрагмента одиночной цепи, который оценивается по автографу, будет соответствовать расстоянию от ее меченого конца до надреза в исходном дуплексе

29.5 На поверхности нуклеосомы структура ДНК неоднородна

Ключевые понятия:

- ДНК обернута вокруг гистонового октамера 1,65 раза.
- В области «обмотки» структура ДНК неоднородна: число пар оснований в витке двойной спирали в середине области больше, чем по краям.

Из-за своего поверхностного расположения, ДНК в нуклеосоме доступна действию некоторых нуклеаз. Особенно информативными оказались реакции нуклеосомной ДНК с нуклеазами, атакующими ДНК по одной из цепей. Ферменты ДНКаза I и ДНКаза II вносят в дуплекс ДНК одноцепочечные надрезы (nicks): разрывают ковалентную связь в фосфодиэфирном остове одной цепи, не трогая связей в цепи напротив. Эффект такого воздействия не заметен, пока ДНК остается двухцепочечной. Но при денатурации (разделении цепи) вместо полноразмерных одиночных цепей можно получить набор коротких фрагментов. Если с самого начала (до обработки ДНКазой) пометить концы ДНК, то концевые фрагменты можно будет идентифицировать авторадиографией, согласно схеме эксперимента на РИС. 29.11. Когда ДНКазы действуют на свободную ДНК в растворе, надрывы (ники) вносятся в нее почти случайным образом. Если же ДНК находится в составе нуклеосом, она тоже может быть надрезана, *но только через регулярные промежутки длины*. Если ДНК с мечеными концами обработать ДНКазой и подвергнуть денатурации, то электрофореграмма будет выглядеть как «лестница»; образец показан на РИС. 29.12.

Интервалы между «ступеньками» такой «лестницы» равняются 10 либо 11 пн. «Лестница» охватывает всю центральную область нуклеосомной ДНК (протяженностью 146 пн). Сайты надрезов пронумерованы от S1 до S13 (где S1 отстоит на ~10 оснований от меченого 5′-конца, S2 — на ~20 оснований и т. д.). Положения сайтов относительно сверхспирали ДНК отмечены на РИС. 29.13.

Не все сайты ДНКаза надрезает с одинаковой частой. Некоторые сайты режутся более эффективно, другие — почти вовсе не подвергаются воздействию. Ферменты

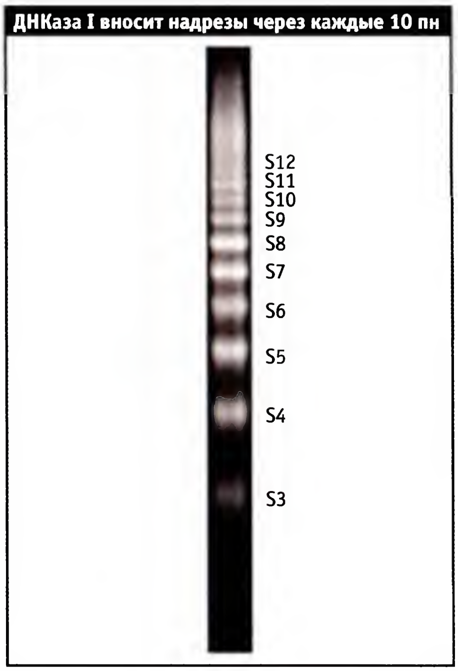


РИС. 29.12 Из автографа видно, что сайты надрезов, вносимых в ДНК центральной частицы, разделены регулярными интервалами. Фотография любезно предоставлена Leonard C. Lutter, Henry Ford Hospital, Detroit, MI

ДНКаза I и ДНКаза II генерируют одинаковые «лестницы», между которыми можно уловить некоторые различия в интенсивности полос. Это говорит о том, что паттерн надрезов представляет собой уникальную серию мишеней, которая определяется организацией (но не последовательностью) ДНК, причем конкретные ферменты отдают незначительные предпочтения определенным сайтам. Точно такой же паттерн расщепления получается при обработке нуклеосомной ДНК гидроксил-радикалом, и это окончательно укрепляет точку зрения, что паттерн расщепления отображает свойства структуры ДНК, а не какую бы то ни было энзиматическую специфичность в отношении ее последовательностей.

Чувствительность нуклеосомной ДНК к нуклеазам объясняется по аналогии с футпринтингом (метод определения участков нуклеиновых кислот, образующих комплексы с белками). Отсутствие реакции по некоторым мишеням можно отнести на счет строения нуклеосомы, при котором определенные участки ДНК пребывают вне доступности.

ДНК центральной частицы нуклеосомы состоит из двух цепей, и (в эксперименте с получением ДНКазной «лестницы») под мечение подпадают оба 5′-конца (или 3′-конца; по одному у каждой цепи; находящиеся на противоположных концах двухцепочечной ДНК центральной области). Таким образом, «лестница» содержит фрагменты обеих цепей. На это обращает внимание рис. 29.11, где два изображенных меченых фрагмента происходят от разных цепей. Вывод из этого обстоятельства такой, что на электрофореграмме каждая полоса («ступенька лестницы») на самом деле соответствует двум фрагментам, происходящим от разных цепей, каждый из которых — продукт надреза ДНК центральной частицы нуклеосомы на *равном* расстоянии от меченого конца.

Как в таком случае следует интерпретировать явные предпочтения, отдаваемые ферментами отдельным сайтам? Одно из мнений на этот счет заключается в том, что траектория ДНК симметрична относительно горизонтальной оси нуклеосомы (как это показано на рис. 29.4). Если, например, ДНКаза I «пропускает ступеньку» из 80 оснований (то есть не генерирует фрагментов такой длины), то это надо понимать так, что в *каждой* из двух цепей позиция, удаленная на 80 оснований от 5′-конца, для фермента недосягаема. Эта точка зрения отражена в альтернативном способе нумерации сайтов, также приведенном на рис. 29.13, где S7 идентифицируется как «сайт 0» и центр симметрии.

Если мобилизовать ДНК на плоской поверхности, то ДНКаза надорвет ее по сайтам, разделенным равными интервалами. На рис. 29.14 показано, что регулярность расположения сайтов, доступных действию ДНКазы, может отображать молекулярную периодичность ДНК (число пар оснований на один виток двойной спирали). Это значит, что расстояние между уязвимыми сайтами равно числу пар азотистых оснований в одном витке двойной спирали ДНК. Согласно проведенным измерениям, среднее значение для свободной двуспиральной ДНК в B-форме составляет 10,5 пн на виток.



РИС. 29.13 Два сверхвитка ДНК центральной частицы нуклеосомы разделены на сегменты по 10 пн. Сайты границ между сегментами можно пронумеровать насквозь, от S1 до S13, либо, выбрав S7 в качестве координаты 0 и центра симметрии, можно пронумеровать их от −7 до +7



РИС. 29.14 В центральной частице нуклеосомы уязвимые сайты ДНК (отмечены только для одной из цепей) распределены с периодичностью, отражающей структуру самой двойной спирали

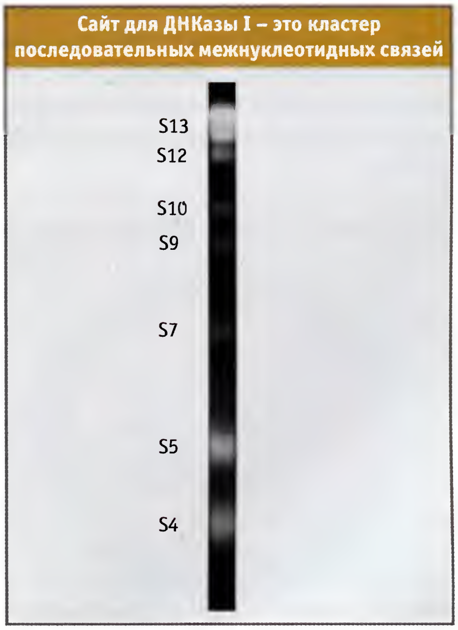


РИС. 29.15 При высоком разрешении автографа видно, что в ДНК центральных частиц каждый сайт для ДНКазы I состоит из нескольких подряд идущих фосфодиэфирных связей. Фотография любезно предоставлена Leonard C. Lutter, Henry Ford Hospital, Detroit, MI

Как же устроены сайты-мишени ДНКазы в нуклеосомной ДНК? На РИС. 29.15 показано, что каждый сайт имеет ненулевую протяженность, точнее, состоит из 3–4 подряд идущих возможных позиций, по одной из которых и вносится надрез. Иными словами, место действия ДНКазы определено с точностью ±2 пн, а сайт-мишень состоит из коротких очередей в 3–4 межнуклеотидных связи — по одной на каждой цепи. Относительные интенсивности полос «лестницы» говорят о предпочтении ферментом одних сайтов другим.

На основании паттерна полос в «лестнице» можно рассчитать усредненный паттерн расположения точек надрезов. По краям центральной области нуклеосомной ДНК сайты с S1 по S4 и с S10 по S13 разнесены на интервалы по 10 пар оснований. В середине центральной области ДНК (и, соответственно, центральной частицы нуклеосомы) сайты с S4 по S10 разделены промежутками по 10,7 пар оснований (здесь подразумеваются *усредненные* интервалы, и значения их длин не обязательно целочисленные).

Разброс значений периодичности расположения сайтов-мишеней ДНКазы в пределах центральной области нуклеосомной ДНК (10,0 — по краям, 10,7 — в середине) говорит о вариабельности молекулярной периодичности ДНК в этой области. В одном витке двойной спирали ДНК в середине области содержится больше пар нуклеотидов, а по краям области — меньше, чем у ДНК, которая плавает в растворе. Молекулярная периодичность ДНК, усредненная по всей нуклеосоме, составляет всего 10,17 пн на виток, и это значительно меньше, чем 10,5 пн на виток у ДНК в растворе.

Анализ кристаллической структуры центральной частицы показал, что ДНК в ней организована как плоская сверхспираль, с 1,65 витков, намотанных на гистоновый октамер. Шаг сверхспирали варьирует и имеет разрывную область значений в середине частицы. Участки с высокой кривизной расположены симметрично, в них находятся сайты ±1 и ±4. Это пары сайтов S6 и S8, S3 и S11, соответственно, именно они менее других чувствительны к ДНКазе I.

Высокое разрешение съемки кристаллической структуры нуклеосомы позволяет детально ознакомиться с искажениями структуры ДНК. В основном сверхспирализация сконцентрирована на центровых 129 пн, что выражается в наличии 1,59 левозакрученных сверхвитков диаметром 80 Å (8 нм, что составляет всего 4 диаметра самой двойной спирали). Края центральной области ДНК вносят очень незначительный вклад в ее общую кривизну.

Центральные 129 пн ДНК пребывают в В-форме, но с избыточной кривизной, необходимой для образования сверхспирали. Большая бороздка изогнута плавно, в то время как малая бороздка имеет резкие изломы. Наличием конформационных неоднородностей, возможно, и объясняется тот факт, что центральная область нуклеосомной ДНК, как правило, не является мишенью для связывания регуляторных белков. Эти белки обычно взаимодействуют с краями центральной области ДНК или с нуклеотидными последовательностями связующей области.

29.6 В нуклеосоме молекулярная периодичность ДНК изменяется

Ключевые понятия:

- Примерно 0,6 избыточного отрицательного сверхвитка ДНК компенсируется изменением ее молекулярной периодичности (числа пн на виток двойной спирали) с 10,5 в растворе до, в среднем, 10,2 на поверхности нуклеосомы; этим объясняется «парадокс порядка зацепления» нуклеосомной ДНК.

Понять структуру нуклеосомной ДНК можно, сравнив сверхспирализованность, рассчитанную исходя из траектории, которой следует ДНК в нуклеосоме, со сверхспирализованностью нуклеосомной ДНК, измеренной альтернативным способом. Предметом большого числа исследований структуры конгломератов нуклеосом был вирус SV40. ДНК SV40 — это кольцевая молекула в 5200 пн, с длиной контура ~1500 нм. В составе вириона, равно как и в ядре инфицированной клетки, такая молекула упакована в виде серии нуклеосом — так называемой минихромосомы.

Длина контура минихромосомы, выделенной обычным способом, составляет 210 нм, что соответствует коэффициенту упаковки ДНК ~7 (по сути, это то же самое, что ~6 в любой нуклеосоме). Поменяв концентрацию соли, можно преобразовать минихромосому в гибкую «нитку бус», причем общий коэффициент упаковки ДНК значительно понизится. Это наглядно демонстрирует способность серий нуклеосом принимать разные формы *in vitro* в зависимости от условий.

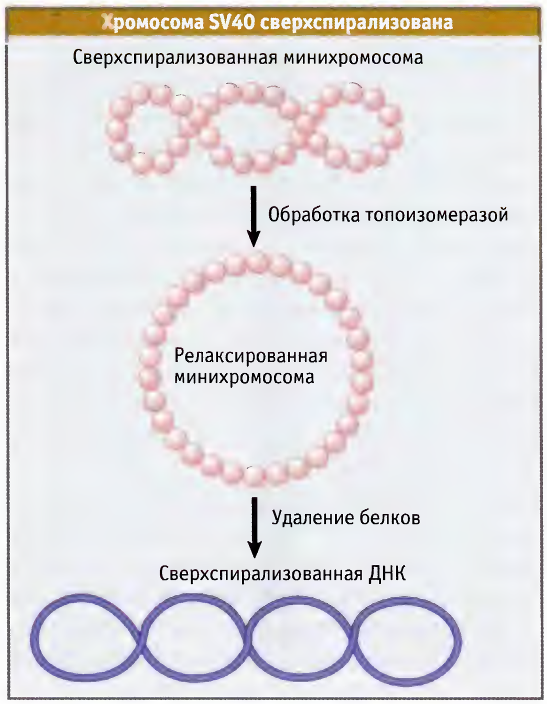


РИС. 29.16 Из минихромосомы SV40 можно убрать сверхвитки; получится релаксированное кольцо. Если отделить ДНК от гистонов, то получится свободная кольцевая сверхспирализованная молекула

Степень сверхспирализованности нуклеосомной ДНК в минихромосоме может быть измерена; принцип изображен на РИС. 29.6. Сначала проводят релаксацию свободных сверхвитков; в результате нуклеосомы образуют кольцевой тяж с плотностью сверхвитков, равной нулю. Затем проводят экстракцию гистоновых октамеров, что предоставляет ДНК свободу принятия непринужденной пространственной конфигурации. В депротеинизированной ДНК каждый сверхвиток, который присутствовал в (релаксированной) минихромосоме, но, вследствие нуклеосомной организации, следовал принудительной траектории, будет присутствовать в явном виде. Теперь суммарное число сверхвитков в ДНК SV40 может быть измерено.

Результат измерения почти совпадает с числом нуклеосом, содержавшихся в минихромосоме, из которых эта ДНК была высвобождена. Такой же результат выдает обратный эксперимент по сборке нуклеосом на сверхспирализованной ДНК SV40 *in vitro*: образование каждой нуклеосомы «убирает» из ДНК ~1 отрицательный сверхвиток.

Таким образом, находясь в составе нуклеосомы, ДНК следует траектории, которая может быть преобразована в ~1 отрицательный сверхвиток. Но траектория, которой ДНК следует, находясь в составе нуклеосомы, со всей очевидностью соответствует −1,67 сверхвитка (см. рис. 29.4) Это расхождение иногда называют «парадоксом порядка зацепления».

Расхождение объясняется разницей между 10,17 пн на виток нуклеосомной ДНК и 10,5 пн на виток свободной ДНК. В нуклеосоме на 200 пн ДНК приходится 200/10,17 = 19,67 витков. После высвобождения этой ДНК из нуклеосомы она содержит уже всего 200/10,5 = 19,0 витков. Редуцированная молекулярная периодичность нуклеосомной ДНК, по сравнению со свободной, поглощает −0,67 сверхвитков двойной спирали. По сути, торсионное напряжение, присутствовавшее в нуклеосомной ДНК, частично расходуется на увеличение числа пн на виток, и только остаток может быть измерен как вклад в сверхспирализованность свободной молекулы.

29.7 Организация гистонового октамера

Ключевые понятия:

- Основой гистонового октамера служит тетрамер $H3_2$–$H4_2$, ассоциированный с двумя димерами H2A–H2B.
- Каждый гистон взаимодействует со своим партнером через обширную поверхность сцепления.
- Все гистоны сердцевины имеют в своей структуре особый мотив — гистоновую складку; N-концевые «хвосты» торчат из нуклеосомы наружу.

До сих пор мы рассматривали нуклеосому с точки зрения организации ДНК на ее поверхности; нуклеосомные белки подробно не рассматривались. Вне сомнения, и белковый аспект организации нуклеосомы немаловажен. Как гистоны взаимодействуют между собой и с ДНК? Формируют ли они октамеры независимо от ДНК или лишь в ее присутствии? Большая часть данных по гистонгистоновым взаимодействиям была получена благодаря способности этих белков к образованию стабильных комплексов, а также на основании экспериментов по внесению в нуклеосому искусственных химических сшивок.

Гистоны сердцевины образуют *in vitro* комплексы двух типов. Гистоны H3 и H4 объединяются в тетрамер ($H3_2$–$H4_2$). Гистоны H2A и H2B могут объединяться в различные комплексы, в частности, димер (H2A–H2B).

Интактные гистоновые октамеры можно получить, либо экстрагировав их из хроматина, либо (что сложнее) позволив гистонам ассоциировать *in vitro* в условиях высоких концентраций соли и самих белков. Октамер может диссоциировать с образованием гексамера гистонов, расставшегося с димером H2A–H2B. Если затем (отдельно от первого) потеряется второй димер H2A–H2B, то останется тетрамер $H3_2$–$H4_2$. Такое течение процесса диссоциации — это аргумент в пользу того, что именно тетрамер $H3_2$–$H4_2$ служит в нуклеосоме опорной конструкцией. Тетрамер может организовывать ДНК *in vitro* в виде частиц, по ряду свойств похожих на центральные частицы нуклеосом.

Опыты с внесением искусственных химических сшивок дополняют сведения о взаимодействии гистонов сердцевины геометрическими подробностями попарного сближения белковых молекул в нуклеосоме. (Сложность интерпретации таких данных связана с тем, что обычно сшивки удается внести лишь в небольшой процент белков, поэтому необходимо соблюдать осторожность, чтобы случайно не завысить степень репрезентативности полученных результатов.) На основании такого рода данных была построена модель организации нуклеосомы. Она представлена на РИС. 29.17.

Рентгеноструктурные исследования, в свою очередь, показали, что по своей форме гистоновый октамер по-

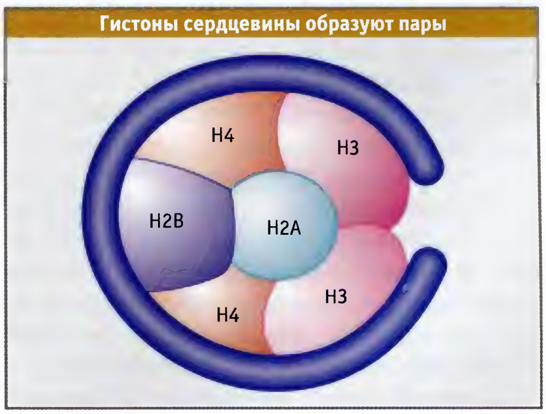


РИС. 29.17 В симметричной модели нуклеосомы тетрамер $H3_2$–$H4_2$ играет роль жесткой основы; он придает сердцевине нуклеосомы ее форму. Из димеров H2A–H2B виден только один; второй димер H2A-H2B находится за плоскостью рисунка

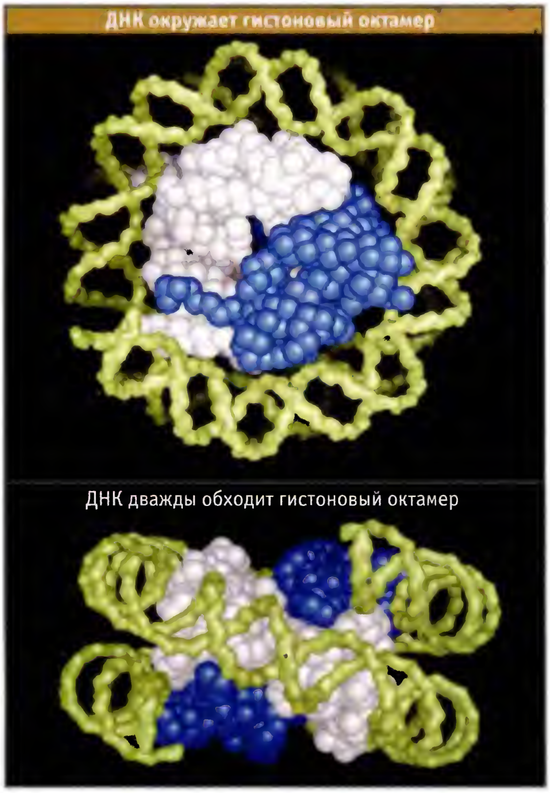


РИС. 29.18 Данные кристаллографического анализа сердцевины нуклеосомы (гистонового октамера) представлены в виде «атомной модели с заполнением пространства». Тетрамер $H3_2$–$H4_2$ здесь имеет белый цвет, а димеры H2A-H2B – синий цвет. Только один из димеров виден сверху, второй находится под ним. «Резьба» для укладки ДНК на виде сверху имеет вид узкой дорожки (в четверть диаметра ДНК); на виде сбоку ее можно увидеть как пучок параллельных линий шириной 20 Å. Фотография любезно предоставлена E. N. Moudrianakis, Johns Hopkins University

хож на центральную частицу нуклеосомы. Это означает, что именно гистонгистоновые взаимодействия определяют строение частицы. Положения отдельных гистонов были отнесены к конкретным областям октамера на основании предпочтений во взаимодействиях и паттернов искусственных сшивок.

Анализ кристаллической структуры, снятой с разрешением 3,1 Å, привел к созданию модели, изображенной на РИС. 29.18. Проследив конфигурации отдельных полипептидов (остовов последовательных пептидных связей одного белка), можно сделать вывод, что гистоны организованы не как обособленные глобулярные белки; каждый из них образует обширное сцепление с партнером по взаимодействию: H3 с H4, H2A с H2B. Хотя тетрамер $H3_2$–$H4_2$ (белый) и димеры H2A–H2B (голубые) на рис. 29.18 хорошо различимы, разделения их на отдельные гистоны в рамках этой модели не проводится.

Вид цилиндрической модели с торца (рис. 29.18, верхняя часть) соответствует схеме на рис. 29.17. Диаметр октамера задается тетрамером $H3_2$–$H4_2$. В торцевом ракурсе он имеет форму подковы. По форме и расположению димеры H2A–H2B вписаны в тетрамер $H3_2$–$H4_2$. Вид сбоку (рис. 29.18, нижняя часть) соответствует схеме на рис. 29.4, но, по сравнению со схемой, функциональные отличия тетрамера $H3_2$–H42 и двух обособленных димеров H2A–H2B здесь очевидны. Весь белковый октамер образует своего рода катушку с намеченной на ней сверхспиральной траекторией, представляющей собой протяженный сайт связывания ДНК. В таком случае ДНК-обмотка в нуклеосоме должна состоять из почти двух полных оборотов (сверхвитков). Модель симметрична относительно оси, перпендикулярной виду сбоку.

В еще больших подробностях расположение гистонов можно представить себе благодаря анализу кристаллической структуры, снятой с разрешением 2,8 Å. Результаты анализа обобщает РИС. 29.19. Верхняя часть

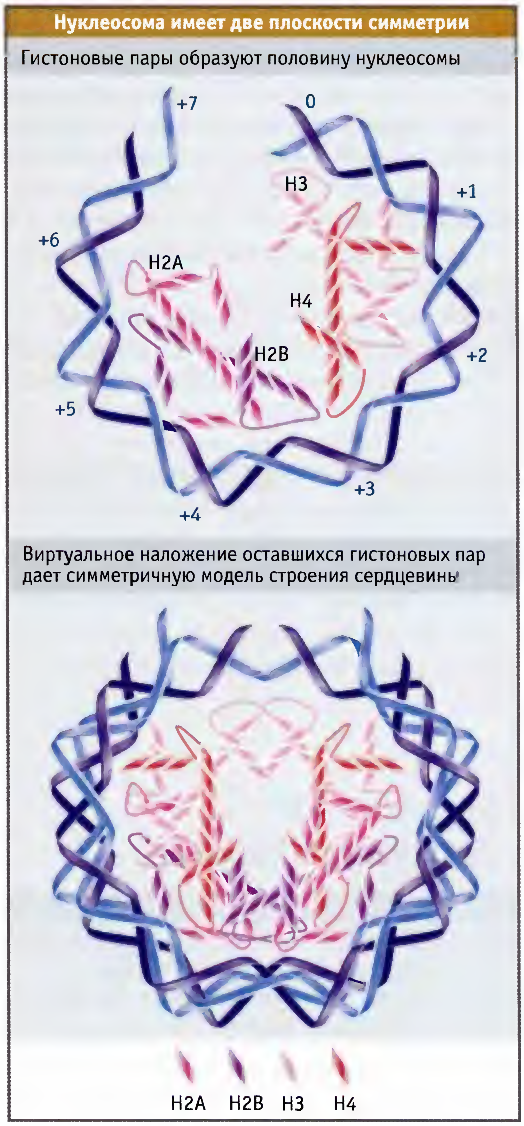


РИС. 29.19 На изображении вверху показаны положения гистоновых пар в половине нуклеосомы; наложение зеркально-симметричной картинки дает структуру целой нуклеосомы, изображенную в нижней части рисунка



РИС. 29.20 Глобулярные тела гистонов находятся в составе гистонового октамера центральной частицы. Положения N-концевых «хвостов», где находится множество регуляторно-значимых сайтов модификации как следует не изучены и, возможно, не постоянны

рисунка показывает по одному гистону каждого типа с их расположением относительно одного оборота ДНК (прокалиброванного номерами от 0 до +7). В структурах всех четырех гистонов сердцевины выявляется одна и та же черта организации белковых доменов: три α-спирали, соединенных двумя петлями. Это сочетание называется **гистоновой складкой**. Гистоновые складки взаимодействуют, с образованием серповидных гетеродимеров; каждый гетеродимер связывает 2,5 оборота двойной спирали ДНК (у H2A-H2B связывается интервал от +3,5 до +6, у $H3_2$–$H4_2$ — интервал от +0,5 до +3 окружности из ДНК, изображенной в верхней части рис. 29.19). Связывание с гистонами затрагивает главным образом фосфодиэфирные остовы ДНК (что согласуется с универсальностью нуклеосом как способа упаковки ДНК вне зависимости от ее нуклеотидной последовательности). Как показано в нижней части рис. 29.19, тетрамер $H3_2$–$H4_2$ образуется благодаря взаимодействию между двумя его H3-субъединицами.



РИС. 29.21 N-концевые «хвосты» гистонов не вписаны в структуру сердцевины и торчат наружу

Каждый из гистонов сердцевины имеет глобулярное «тело», участвующее в формировании центрального белкового массива нуклеосомы. Каждый гистон имеет также гибкий N-концевой «хвост», содержащий сайты для модификаций, которые могут быть важны для функционирования хроматина. Положения «хвостов», на которые приходится примерно четверть массы белков, установлены не с такой определенностью, какая продемонстрирована на РИС. 19.20. Все же «хвосты» H3 и H2B можно увидеть выходящими из нуклеосомы между оборотами сверхспирали ДНК и торчащими наружу, как показано на РИС. 19.21. Если «хвосты» гистонов присоединить к ДНК под действием УФ-облучения, то цельные нуклеосомы дают больше продуктов такой реакции, чем центральные частицы. Это может означать, что «хвосты» имеют контакты с ДНК связующей области. «Хвост» H4, судя по всему, контактирует с димером H2A–H2B смежной нуклеосомы; это может быть важной чертой интегральной структуры хроматина.

29.8 Укладка нуклеосом в волокнах хроматина

Ключевые понятия:

- 30-нм волокна хроматина раскручиваются в цепочку нуклеосом толщиной 10 нм.
- 30-нм волокно устроено как спираль высшего порядка, на один виток которой приходится шесть нуклеосом.
- Для формирования 30-нм волокна требуется гистон H1.

При изучении хроматина под электронным микроскопом можно разглядеть волокна двух типов, с диаметрами 10 нм и 30 нм. Диаметр.указан приблизительно, у 30-нанометрового волокна он, вообще-то варьирует в пределах ~25–30 нм.

10-нанометровое волокно (10-нм волокно) представляет собой непрерывную череду нуклеосом. На самом деле, время от времени оно непрерывно перетекает в более растянутые участки, где нуклеосомы похожи на бусины, надетые на нитку; фотография такого участка хроматина представлена на РИС. 29.22. Увидеть 10-нанометровое волокно *in vitro* можно в растворе с низкой ионной силой и при отсутствии гистона H1. Это означает, что его существование — это функция самих нуклеосом. Изобразить такое волокно можно, как это сделано на РИС. 29. 23, в виде непрерывной серии нуклеосом. Существует ли такая структура сама по себе *in vivo*, неясно; возможно, она представляет собой артефакт — следствие частичного самопроизвольного роспуска структуры хроматина в процессе его экстракции.

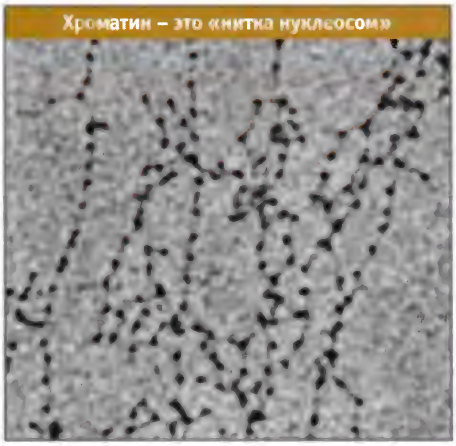


РИС. 29.22 Когда десятинанометровое волокно находится в частично расплетенном состоянии, то видно, что оно представляет собой череду нуклеосом. Фотография любезно предоставлена Barbara Hamkalo, University of California, Irvine



РИС. 29.23 Десятинанометровое волокно – это непрерывная череда нуклеосом

Наблюдая хроматин *in vitro* в растворах с большей ионной силой, можно увидеть **30-нанометровое волокно**; пример приведен на РИС. 19.24. Можно даже разглядеть, что такое волокно имеет спиральную структуру, на каждый оборот которой приходится ~6 нуклеосом. Коэффициент упаковки ДНК в 30-нанометровом волокне равняется 40 (это значит, что каждый микрометр волокна — вдоль оси — содержит 40 мкм ДНК). Для сохранности волокна необходимо присутствие в нем гистона H1. 30-нанометровое волокно — это основная составляющая как интерфазного хроматина, так и митотических хромосом.

Наиболее вероятная схема укладки нуклеосом в 30-нанометровом волокне — это соленоид, или спиральная обмотка центральной полости 10-нанометровым волокном. Две основных формы такого соленоида — однорядный (single-start; из одиночного линейного массива нуклеосом) и двурядный (two-start; он скручен из двух тяжей нуклеосом). На РИС. 29.25 показана модель двурядного соленоида, в котором один тяж нуклеосом уложен в два ряда в виде ленты, которая и служит обмоткой соленоида; модель основана на данных последних экспериментов по внесению искусственных сшивок (crosslinking). Согласно этой модели, в двурядном соленоиде ДНК перебегает с одного 10-нанометрового тяжа на другой. Это подтверждается анализом кристаллической структуры тетрануклеосомного комплекса.

30-нанометровое и 10-нанометровое волокна могут обратимо превращаться одно в другое при изменениях ионной силы раствора, в котором находится материал.

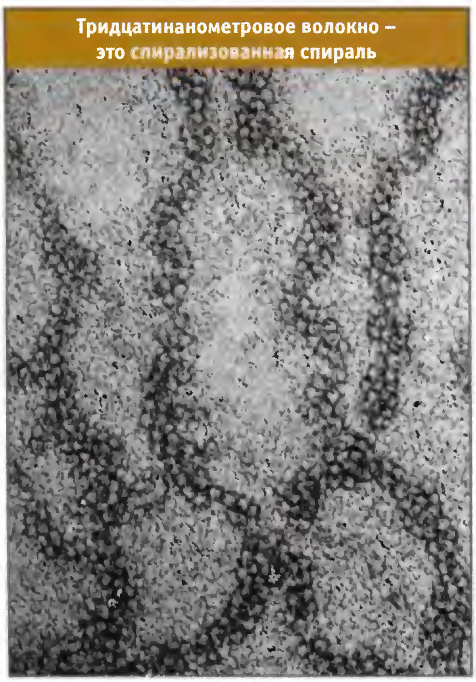


РИС. 29.24 Тридцатинанометровое волокно имеет спиралевидную структуру. Фотография любезно предоставлена Barbara Hamkalo, University of California, Irvine

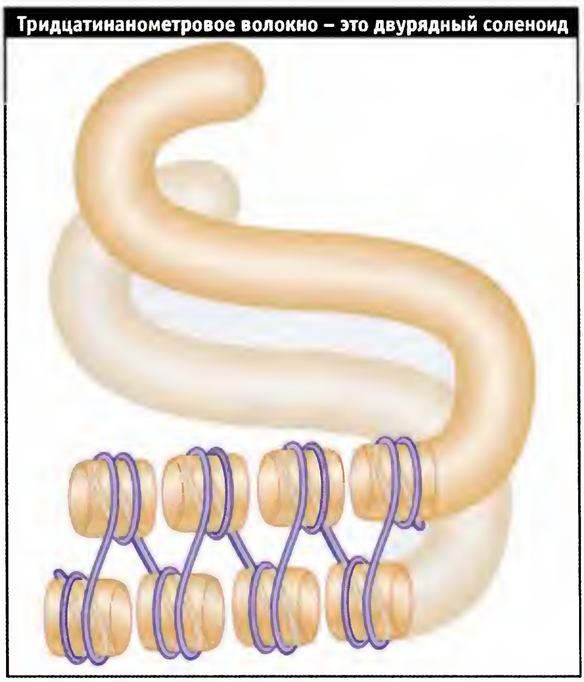


РИС. 29.25 Тридцатинанометровое волокно состоит из двух параллельных рядов нуклеосом, уложенных в спираль

По всей видимости, линейный массив нуклеосом — 10-нанометровое волокно — свивается в фибриллярную структуру диаметром 30 нм — при увеличении ионной силы, в присутствии гистона H1.

Присутствие H1 необходимо для формирования 30-нанометрового волокна, однако информация о местоположении гистона в самом волокне довольно противоречива. Относительная легкость, с которой H1 удается экстрагировать из хроматина, указывает, что он находится на наружной поверхности волокна. Но существуют данные дифракционного анализа, которые свидетельствуют о расположении H1 внутри волокна; а также тот факт, что H1 не так легко выявляется в 30-нанометровых волокнах, как в 10-нанометровых, которые хорошо удерживают этот гистон.

Как устроен переход от 30-нанометрового волокна к тем особым структурам, которые видны в митотических хромосомах? Существует ли какая-либо высокоуровневая специфичность распределения хроматина в интерфазном ядре? Заданы ли взаимные положения конкретных участков 30-нанометрового волокна хоть сколько-нибудь жестко, или их распределение в интерфазном ядре носит случайный характер?



РИС. 29.26 Реплицированная ДНК сразу же включается в нуклеосомы. Фотография любезно предоставлена Steven L. McKnight, UT Southwestern Medical Center at Dallas

29.9 Воспроизводство хроматина требует сборки нуклеосом

Ключевые понятия:

- Во время репликации гистоновые октамеры из расформированных нуклеосом матричного дуплекса разбираются на димеры H2A–H2B и тетрамеры $H3_2$–$H4_2$.
- Сборка нуклеосом идет разными путями в зависимости от того, происходит ли она во время репликации ДНК или независимо от нее.
- Сборка нуклеосом требует участия вспомогательных белков.
- Сборочный белок CAF-1 контактирует с PCNA-субъединицей реплисомы; он требуется для депонирования тетрамеров $H3_2$–$H4_2$ и их выкладывания на дочерние дуплексы после репликации.
- При независимой от репликации сборке нуклеосом используется другой сборочный белок, а также альтернативная разновидность гистона H3.

При репликации цепи матричного дуплекса ДНК разделяются. Этому неизбежно должно сопутствовать расформирование нуклеосом, в состав которых они входили. Анализировать структурную динамику конкретного реплицируемого участка очень трудно — из-за быстротечности события репликации. У репликативной вилки структура особенная, ее ДНК более устойчива к микрококковой нуклеазе, чем нуклеосомная ДНК остального хроматина, и при расщеплении дает паттерн полос (длин фрагментов), отличающийся от аналогичных паттернов нуклеосомной ДНК. Эта особенная структура ДНК ограничивается ближайшей окрестностью репликативной вилки. Такое впечатление, что большой белковый комплекс, реплицирующий ДНК, несовместим с нуклеосомами — но они снова формируются, сразу после прохождения реплицирующего комплекса, *только уже на дочерних дуплексах*.

В воспроизводстве хроматина отсутствует детектируемое временное окно, когда ДНК была бы свободна от гистонов. Как только участок ДНК реплицировался, на обоих дочерних дуплексах быстро возникают нуклеосомы; это хорошо видно на электронной микрофотографии РИС. 29.26.

Итак, данные биохимического анализа и электронной микроскопии сходятся на том, что нуклеосомы пребывают расформированными лишь в пределах короткого участка ДНК — окрестности репликативной вилки. Продвигаясь, вилка сокрушает нуклеосомы на своем пути, но, когда вилка проходит, они очень быстро возникают на дочерних дуплексах. На самом деле, сборка нуклеосом имеет самое прямое отношение к аппарату репликации ДНК.

Каким образом гистоны связываются с ДНК, чтобы сформировалась нуклеосома? Формируют ли гистоны октамер *предварительно* (и ДНК на него наматывается), или сборка октамера проходит прямо на ДНК? На РИС. 29.27 показаны два пути, которыми может идти сборка нуклеосом *in vitro*; выбор пути зависит от условий реакционной среды. В соответствии с одним из путей, с ДНК связывается октамер, сформированный заранее. Второй путь — это связывание с ДНК тетрамера $H3_2$–$H4_2$ с последующим присоединением к комплексу двух димеров H2A–H2B. Каждый из путей соответствует определенным событиям *in vivo*. Первый путь может быть соотнесен с таким свойством хроматина, как податливость к ремоделированию, и отображает перемещение гистоновых октамеров вдоль ДНК (см. разд. 30.3 «Ремоделирование хроматина — это активный процесс»). Второй путь отображает способ сборки нуклеосом, используемый при репликации.

Ассоциации гистонов с ДНК помогает ряд белков (accessory proteins). Их первичную идентификацию проводили, используя ядерные экстракты, способные

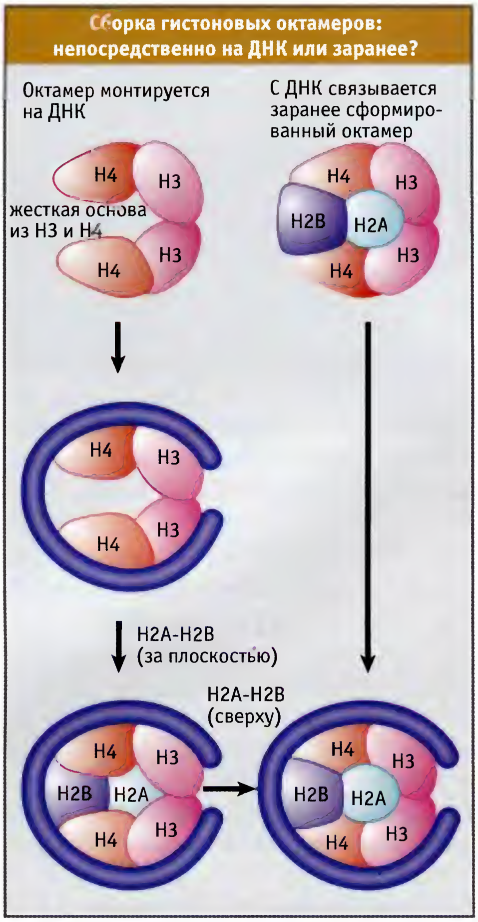


РИС. 29.27 В условиях *in vitro* нуклеосомы образуются от взаимодействия ДНК сразу с целым гистоновым октамером. В других экспериментах ДНК взаимодействовала с тетрамером $H3_2$–$H4_2$ до его дополнения двумя димерами H2A–H2B

собирать нуклеосомы из гистонов и экзогенной ДНК. Вспомогательные белки могут работать «молекулярными шаперонами», то есть связываться с гистонами для того, чтобы помещать их на ДНК — либо как отдельные белки, либо в виде комплексов ($H3_2$–$H4_2$ или H2A–H2B). Причем делать это нужно в контролируемом режиме, поскольку гистоны, будучи основными (положительно заряженными) белками, обладают высоким общим сродством к ДНК. *Взаимодействия со вспомогательными белками позволяют гистонам формировать нуклеосомы, не вступая в альтернативные комплексы, которые являются результатом случайного связывания гистонов с ДНК.*

Попытки продуцировать нуклеосомы *in vitro* начинались с изучения процесса сборки при участии свободной ДНК и гистонов. Однако в условиях *in vivo* серии нуклеосом образуются *de novo* только тогда, когда идет репликация. Система, которая была разработана с учетом этого обстоятельства, имитирует репликацию клеточной ДНК: в ней используются экстракты из клеток человека, реплицирующие ДНК SV40, а продукты репликации, структурирующие как хроматин. Используя эту систему, удалось выяснить, что для реакции сборки нуклеосом на реплицирующейся ДНК требуется вспомогательный фактор, получивший название CAF-1 (chromatin assembly factor; «фактор сборки хроматина»). Это белок, состоящий из более чем 5 субъединиц, суммарной массой 238 кДа. Его «призывает» в репликативную вилку PCNA (proliferating cell nuclear antigen; «ядерный антиген пролиферирующих клеток») — белок, который является частью реплисомы и служит фактором процессивности ДНК-полимеразы. Контакт между PCNA и CAF-1 обеспечивает сопряжение между процессами репликации и сборки нуклеосом, так что сборка нуклеосом происходит на продукте репликации сразу, по мере его возникновения.

Действие CAF-1 стехиометрично (то есть его количество пропорционально размерам сферы его действия — некоторой области ДНК). Связываясь с гистонами, CAF-1 организует сборку сердцевины будущей нуклеосомы; скорее всего, сначала собирается тетрамер $H3_2$–$H4_2$, к которому затем добавляются димеры H2A–H2B. На образование одной нуклеосомы уходит 200 пн ДНК, причем в данном случае это разделение на правильные интервалы происходит в отсутствие гистона H1.

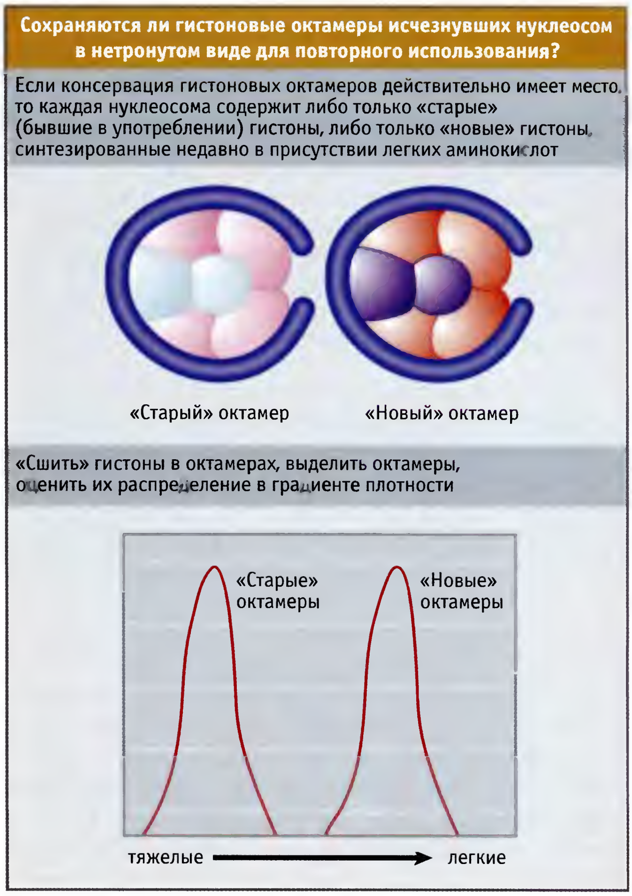


РИС. 29.28 Если гистоновые октамеры сохраняются, то в условиях данного эксперимента их распределение по массе должно иметь два дискретных пика

При воспроизводстве хроматина начало двум дочерним дуплексам дает ДНК, которая *уже структурирована в виде нуклеосом*. Что происходит с этими нуклеосомами, когда они оказываются застигнутыми репликацией? Может быть, гистоновые октамеры диссоциируют на свободные гистоны, которые впоследствии реутилизируются? Или же они остаются до какой-то степени ассоциированными друг с другом? Целостность октамера можно проверить, зафиксировав положения гистонов искусственными сшивками. Два следующих рисунка посвящены сравнению возможных исходов одного эксперимента. Схема эксперимента такова: клетки выращивают в присутствии тяжелых аминокислот, чтобы гистоны, существовавшие до репликации, можно было отличить. Затем, заменив тяжелые аминокислоты легкими, проводят репликацию. Сразу же после репликации гистоновые октамеры объядиняют искусственными сшивками и центрифугируют в градиенте плотности. На РИС. 29.28 показано, что, если бы на время репликации гистоновые октамеры консервировались, то, в конце эксперимента, их появления следовало бы ожидать в положении, соответствующем высокой плотности, в то время как новые октамеры занимали бы в градиенте положение, соответствующее низкой плотности. Этого как раз и не происходит. Максимальную плотность имеет лишь малая часть материала, и это указывает на отсутствие консервации гистоновых октамеров. Как показано на РИС. 29.29, распределение плотностей октамеров разбросано в интервале между высоким и низким значениями. Это объясняется смешением старых гистонов, высвобождающихся из сердцевин расформированных нуклеосом с новосинтезированными гистонами: формируются новые нуклеосомы, вбирающие в себя реплицированный материал.

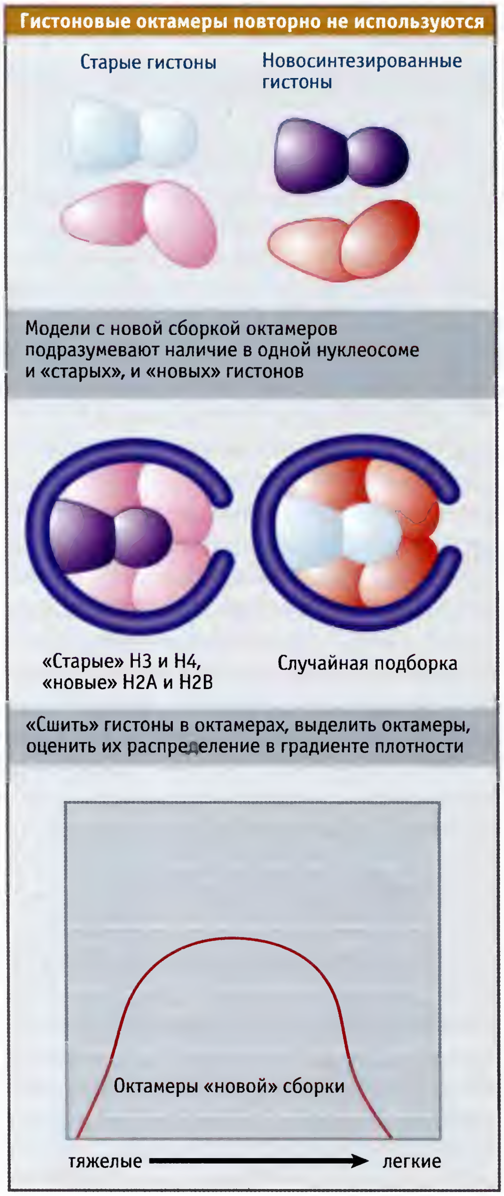


РИС. 29.29 Массы гистоновых октамеров равномерно распределены в диапазоне возможных значений; это говорит об отсутствии консервации октамеров между их использованиями в качестве сердцевин нуклеосом

Режим, в котором происходит разборка и сборка нуклеосом при репликации, довольно сложно охарактеризовать в деталях; рабочая модель изображена на РИС. 29.30. Репликативная вилка смещает гистоновые октамеры; они диссоциируют на тетрамеры $H3_2$–$H4_2$ и димеры H2A–H2B. Эти старые тетрамеры и димеры вливаются в общий пул тетрамеров и димеров, часть которых собрана из новосинтезированных гистонов. Сборка нуклеосом происходит за репликативной вилкой, на расстоянии ~600 пн. Сборка инициируется связыванием тетрамеров $H3_2$–$H4_2$ с дочерними дуплексами; CAF-1 помогает тетрамерам связаться с ДНК. Когда с находящимся на ДНК тетрамером $H3_2$–$H4_2$ свяжутся два димера H2A–H2B, получится укомплектованный октамер. Подбор субъединиц каждого октамера происходит случайно, безотносительно их новизны. Если бы старый тетрамер был способен ненадолго удержаться на одиночной цепи ДНК, то это увеличило бы его шансы остаться с ведущей цепью и быть использованным заново. Возможно, похожим образом нуклеосомы разбираются и собираются и при транскрипции (см. разд. 29.11 «Организованы ли транскрибируемые гены в нуклеосомы?»).

В продолжение S-фазы клеточного цикла эукариотической клетки, когда происходит репликация ДНК, удвоение количества хроматина требует синтеза гистонов в количестве, достаточном для нуклеосомной упаковки целого генома. По сути, гистонов в S-фазе должно быть синтезировано столько, сколько их уже присутствует в нуклеосомах на момент начала репликации. Синтез гистоновых мРНК контролируется как составляющая клеточного цикла; в S-фазе транскрипция гистоновых генов происходит интенсивно. Процесс сборки хроматина из смеси равных количеств «старых» и «новых» гистонов, который имеет место в S-фазе при репликации генома, обозначают как сопряженный с репликацией (replication-coupled pathway, RC pathway).

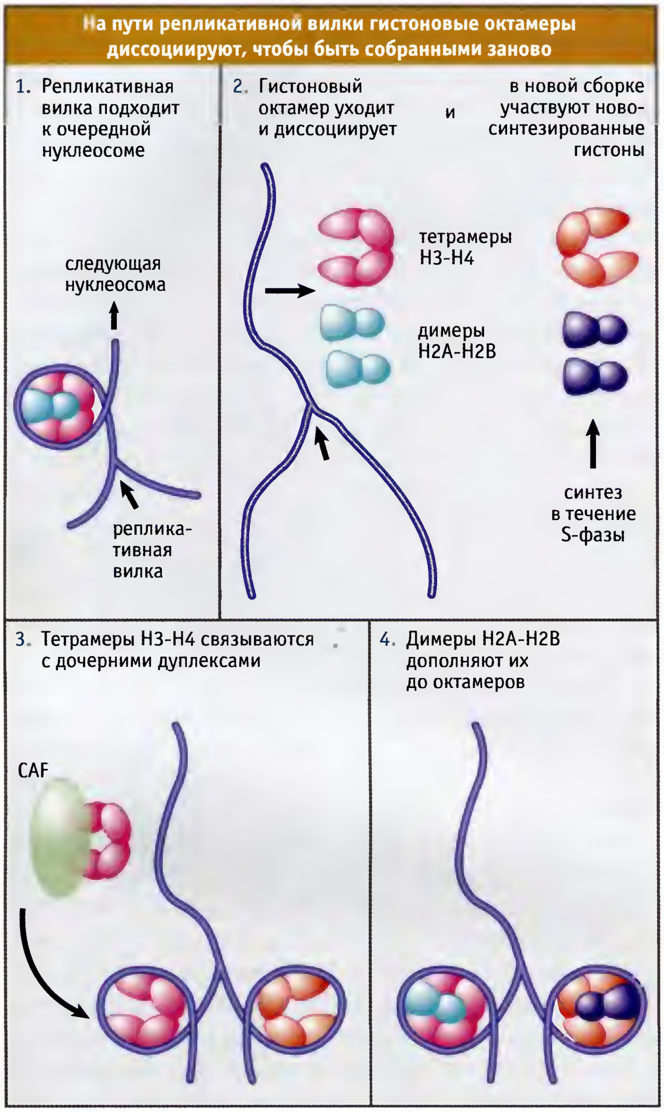


РИС. 29.30 **Репликативная вилка, продвигаясь вдоль ДНК, смещает с нее гистоновые октамеры. Они диссоциируют на тетрамеры $H3_2$–$H4_2$ и димеры H2A-H2B. Тем временем из новосинтезированных гистонов собираются новые тетрамеры $H3_2$–$H4_2$ и димеры H2A–H2B. При участии фактора CAF-1 «старые» и «новые» тетрамеры и димеры (в случайных сочетаниях) собираются в октамеры новых нуклеосом. Это происходит непосредственно за репликативной вилкой**

Для сборки нуклеосом в других фазах клеточного цикла, когда ДНК не синтезируется, существует другой — репликационно-независимый процесс (replication-independent pathway, RI pathway). Необходимость в нем возникает при повреждениях ДНК или при смещении нуклеосом во время транскрипции. Этот процесс сборки неизбежно должен отличаться от процесса сборки, сопряженного с репликацией: хотя бы потому, что он не может быть координирован клеточным аппаратом репликации. Одна из интересных особенностей репликационно-независимой сборки нуклеосом заключается в использовании альтернативных разновидностей некоторых гистонов.

Гистон H3.3, являясь разновидностью высококонсервативного H3, отличается от него по четырем аминокислотным положениям. В дифференцирующихся клетках, где не происходит репликации, H3.3 медленно замещает H3. Это происходит в результате сборки новых гистоновых октамеров на замену тем, которые по какой-либо причине оказались смещенными с ДНК. Механизмы, гарантирующие употребление именно H3.3 в сборке новых октамеров в клетках, которые дифференцируются или дифференцированы, были исследованы у *Tetrahymena* и *Drosophila*; у этих объектов они оказались различными.

У простейшего *Tetrahymena* употребление гистона определяется исключительно его доступностью. Гистон H3 синтезируется, пока клетка пребывает в цикле деления; в клетках, покинувших цикл и прекративших реплицировать свою ДНК, начинает синтезироваться вариант H3.3. У *Drosophila*, напротив, существует активный механизм контроля, который гарантирует использование гистона H3.3. Независимая от репликации сборка новых нуклеосом происходит в сайтах транскрипции. Эти новые нуклеосомы, по всей видимости, идут на замену тем, которые были смещены РНК-полимеразой. Разница между H3 и H3.3, обусловленная различиями в их аминокислотных последовательностях, приводит к тому, что H3 специфически отстраняется от участия в построении октамеров. Сборка нуклеосом, сопряженная с репликацией, у той же *Drosophila*, напротив, задействует оба варианта, H3 и H3.3. На момент репликации H3.3 доступен для употребления в количествах, гораздо меньших, чем H3, и потому оказывается включенным только в небольшой процент новых нуклеосом.

Белок CAF-1, вероятнее всего, не участвует в независимом от репликации процессе сборки. (Более того, у некоторых организмов, таких как дрожжи или *Arabidopsis*, отсутствие его гена не критично, что наводит на мысль о существовании у этих организмов альтернативного механизма сборки нуклеосом при репликации.) Белок, который причастен к независимой от репликации сборке, называют HIRA. Если взять систему сборки нуклеосом *in vitro* и исключить из нее HIRA, то образование нуклеосом будет подавлено только на нереплицированной ДНК, а на реплицирующейся ДНК продолжится без изменений. Это еще раз подтверждает, что процессы сборки подчинены разным механизмам. Белок HIRA работает шапероном, содействующий инкорпорации гистонов в нуклеосомы. Возможно, его участие требуется всегда, когда сборка нуклеосом идет на ДНК вне состояния репликации: например, когда, при деконденсации хроматина в ядре сперматозоида, протамины заменяются гистонами, чтобы образовался хроматин, способный к репликации в составе оплодотворенной яйцеклетки.

Сборка нуклеосом с использованием альтернатив гистону H3 также происходит в центромерах (см. разд. 31.3 «Существование гетерохроматина обусловлено взаимодействиями гистонов с другими белками»). Центромерная ДНК реплицируется в ранней S-фазе (в отличие от окружающих ее последовательностей из состава

гетерохроматина, которые реплицируются позднее; см. разд. 15.7 «В геномных последовательностях дрожжей можно выявить ориджины репликации»). На центромерной ДНК включение H3 в нуклеосомы подавлено; вместо него включается белок, который у высших эукариот известен как CENP-A (у *Drosophila* его называют Cid, у дрожжей — Cse4). Механизм, контролирующий этот процесс, репликационно-независим. Доминирующая в S-фазе сборка по канонической, сопряженной с репликацией схеме, очевидно, подавляется — на короткий период времени, пока реплицируется центромерная ДНК.

29.10 Специфично ли расположение нуклеосом относительно ДНК?

Ключевые понятия:

- Нуклеосомы могут формироваться в определенных положениях под влиянием локальной структуры ДНК либо под влиянием белков, взаимодействующих со специфическими последовательностями.
- Самая распространенная причина позиционирования нуклеосом — это разграничение ДНК связанными с ней белками.
- Позиционирование может определять, какие последовательности ДНК в серии нуклеосом окажутся в связующих областях, и какой стороной наружу будет смотреть ДНК с поверхности гистоновых октамеров.

Нуклеосомы, как мы уже знаем, могут быть воссозданы *in vitro* при участии ДНК с произвольной последовательностью. Но это не означает, что их образование *in vivo* никак не обусловлено первичной структурой ДНК. Насколько однозначно задана топография нуклеосом на последовательностях ДНК в хроматине *in vivo*? Или нуклеосомы расположены на ДНК случайно, так, что конкретная последовательность может занять любое положение; например, у одной копии генома оказаться в центральной области нуклеосомы, а у другой — в связующей области?

Чтобы ответить на этот вопрос, нужно, для начала, взять конкретную последовательность ДНК и определить, какое множество положений может занимать в составе нуклеосомы определенная точка этой последовательности. Принцип используемой для этого экспериментальной процедуры изображен на РИС. 29.31.

Предположим, что последовательность ДНК переводится в состав серии нуклеосом всегда в одной и той же конфигурации. То есть любой сайт ДНК всегда занимает в нуклеосоме определенное положение. Такой тип организации называют **позиционированием нуклеосом**, или их фазированием (nucleosome positioning, nucleosome phasing). В сериях позиционированных нуклеосом связующие области образуют уникальный паттерн.

Иными словами, при обработке однотипных серий позиционированных нуклеосом микрококковой нуклеазой их ДНК будет расчленяться на *специфические последовательности*. Если эти специфические последовательности обработать эндонуклеазой рестрикции, узнающей в одной из них уникальный сайт, то окажется, что этот сайт всегда будет находиться в одном и том же положении относительно концов фрагмента, который его содержит.

Продукты двойного расщепления (микрококковой нуклеазой и эндонуклеазой рестрикции) разделяют гель-электрофорезом. Меченую гибридизационную пробу, по своей нуклеотидной последовательности соответствующую одному из флангов сайта рестрикции, используют для визуализации содержащего ее фрагмента в продукте двойного расщепления. Методика получила название **непрямого концевого мечения**.

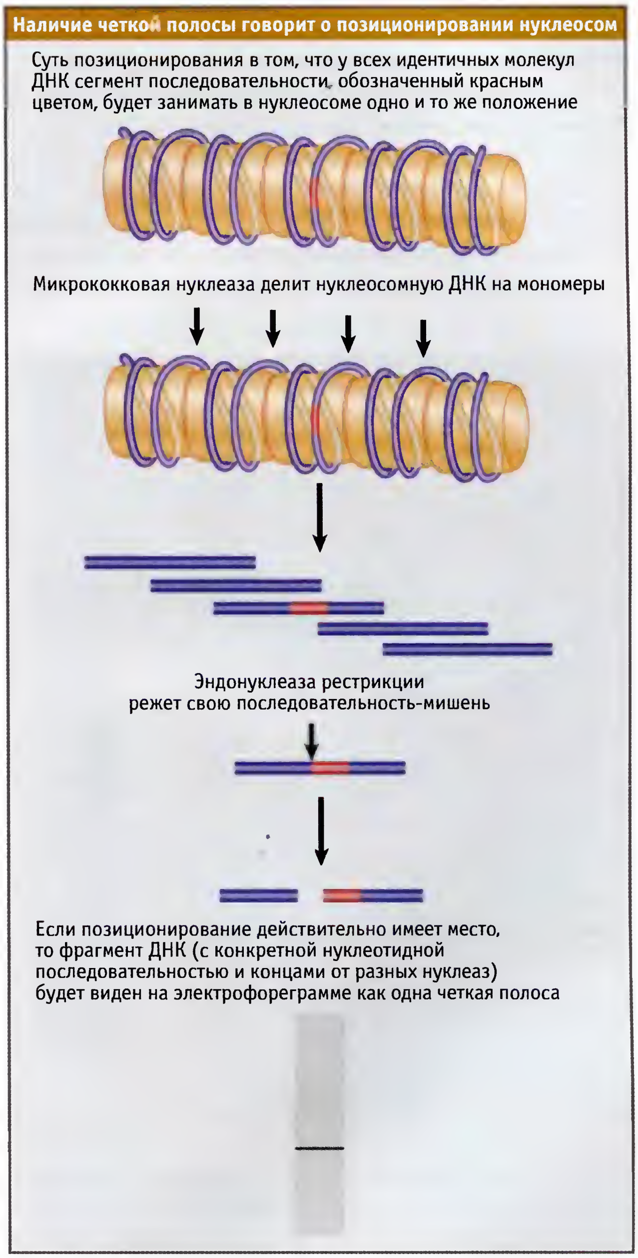


РИС. 29.31 Если нуклеосомы позиционированы относительно нуклеотидной последовательности ДНК, то сайты рестрикции находятся на фиксированном расстоянии от связующих сегментов ДНК, чувствительных к микрококковой нуклеазе

Повернув рассуждение вспять, интерпретируем результаты. Одна четкая полоса говорит о том, что во всех клетках, пошедших на приготовление образца, данный сайт рестрикции занимает одинаковое положение относительно содержащей его нуклеосомы. Одна четкая полоса означает, что для данного участка генома характерно уникальное разбиение последовательности ДНК на нуклеосомные фрагменты.

В противном случае, когда уникального разбиения не существует, границы фрагментов, на которые геномная последовательность разбивается при сборке нуклеосом, статистически «размыты». Поэтому положение сайта рестрикции будет непостоянным относительно нуклеосомы, в которой он находится. Для каждой копии генома связующие области серии нуклеосом, по которым идет расщепление микрококковой нуклеазой, накладываются на последовательность ДНК случайным образом. Каждый раз (в каждой копии генома, подпавшей под двойное расщепление) сайт рестрикции будет занимать любое из возможных положений относительно концов мономера нуклеосомной ДНК. На РИС. 29.32 показано, что в этом случае непрямое концевое мечение дает шмер (от *англ.* smear — размытость границ дорожки в геле) на интервале от минимальной детектируемой длины (~20 оснований) до длины мономера нуклеосомной ДНК (~200 оснований).

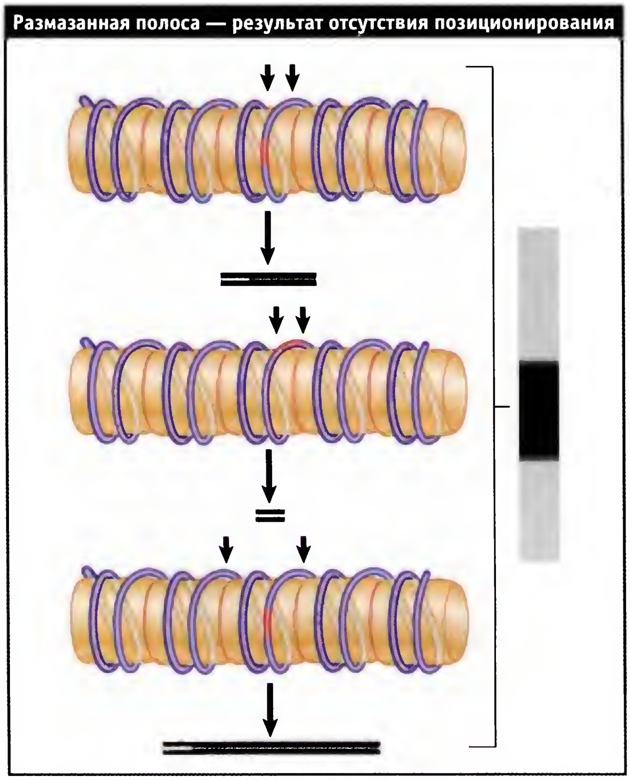


РИС. 29.32 Если нуклеосомы не позиционированы, конкретный сайт рестрикции может занять в нуклеосоме любое положение, и когда эндонуклеаза рестрикции разрежет свою конкретную мишень (обозначенную красным цветом), а микрококковая нуклеаза — связующий сегмент ДНК (обозначенный зеленым цветом), получится фрагмент случайной длины. Электрофореграмма показывает шмер из фрагментов разной длины в диапазоне от 0 до ~200 пн единичной длины нуклеосомной ДНК

До сих пор обсуждение экспериментов с микрококковой нуклеазой строилось на том, что этот фермент разрезает ДНК по доступным ему связующим областям нуклеосомной ДНК вне какой-либо специфичности относительно нуклеотидных последовательностей этих областей. На самом деле, кое-какая специфичность этому ферменту все-таки свойственна; он предпочитает А-Т-богатые последовательности. Так, если непрямое концевое мечение выявило на электрофореграмме четкую полосу, эта полоса не обязательно соответствует расстоянию от сайта рестрикции до *ближайшей* связующей области: она может соответствовать расстоянию до *другой* связующей области, которую предпочла нуклеаза! Из тех же соображений, появление нескольких полос может отображать незавершенность разреза нуклеазой, но может быть истолковано и как наличие нескольких дискретных вариантов позиционирования нуклеосом.

Чтобы корректно интерпретировать такого рода результаты, ставится контрольный эксперимент, в котором, вместо хроматина, микрококковой нуклеазой обрабатывается депротеинизированная ДНК. Все предпочтения микрококковой нуклеазы в пределах изучаемой области ДНК при этом становятся явными: они отображаются в виде паттерна полос, который можно сравнить с паттерном полос, полученном на образце хроматина.

Различия между паттернами полос в контрольном опыте (с депротеинизированной ДНК) и в опыте с хроматином — свидетельство событий позиционирования нуклеосом. Некоторые полосы, присутствующие в контрольном образце, исчезают у продукта расщепления нуклеосом в связи с недоступностью для нуклеазы ряда сайтов (в областях нуклеосом). Некоторые полосы, наоборот, появятся у образца с нуклеосомной организацией, так как новые сайты получат приоритетную доступность для расщепления.

Позиционирование нуклеосом может быть выполнено одним из двух способов:

- Собственно позиционирование: *каждая нуклеосома специфично связывает конкретный участок последовательности ДНК.* Такой способ позиционирования меняет видение нуклеосом как структурных единиц хроматина, возникающих при контакте между гистоновым октамером и любой последовательностью ДНК.
- Наведенное позиционирование: *в серии нуклеосом, покрывающей некоторый участок генома, сборка первой нуклеосомы чаще всего происходит в определенном сайте.* Сама возможность предпочесть какой-либо сайт в качестве стартовой точки позиционирования участка генома проистекает от наличия участка, на котором образование нуклеосом исключается. Положение, по которому может возникнуть нуклеосома, первая в серии, определяется *границей* свободного участка. За пределами

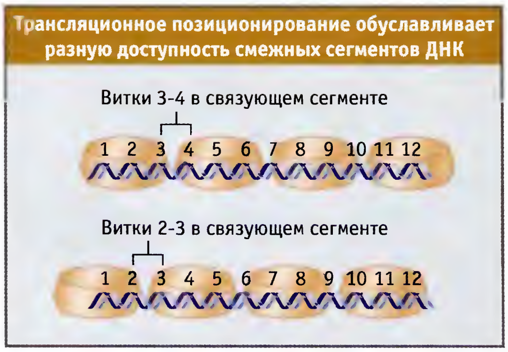


РИС. 29.33 Под трансляционным позиционированием понимают положение линейной нуклеотидной последовательности относительно гистоновых октамеров. Каждый сдвиг ДНК на 10 пн меняет в ней наиболее доступные связующие сегменты. При этом надо иметь в виду, что при контакте ДНК с гистонами сердцевины одна из сторон двойной спирали остается доступной для факторов и нуклеаз, и при сдвиге октамеров на 10 пн эта сторона не меняется. Для простоты схемы нуклеосомная ДНК изображена без сверхвитков

этого участка образование нуклеосом может начаться вновь — последовательно, с фиксированной длиной нуклеосомных фрагментов ДНК.

В настоящее время установлено, что расстановка гистоновых октамеров на ДНК происходит не случайно, но в зависимости от последовательности ДНК. Иногда специфический паттерн позиционирования диктуется особенностями первичной структуры ДНК на конкретном участке. В других случаях специфический паттерн позиционирования является наведенным; он определяется взаимодействием других белков с ДНК и/или гистонами.

Какие структурные свойства самой ДНК влияют на размещение гистоновых октамеров? ДНК с конкретной нуклеотидной последовательностью проявляет склонность к самопроизвольному изгибанию в ту или иную сторону; например, А-Т-богатые участки на поверхности нуклеосомы расположены так, что малая бороздка обращена к октамеру, а если на поверхности нуклеосомы находится G-C-богатый участок, то его малая бороздка смотрит наружу. Длинные очереди dA-dT (>8 пн) избегают позиционирования в центральном сверхвитке нуклеосомы. Просуммировать все сколько-нибудь существенные структурные эффекты с тем, чтобы максимально точно предсказать расположение конкретной последовательности ДНК относительно массива нуклеосом, пока невозможно. Иногда в геноме встречаются экстремальные последовательности ДНК, несовместимые с нуклеосомной организацией. Неблагоприятная конформация ДНК исключает возникновение нуклеосом на таких участках — с эффектом установления границы для наведенного позиционирования нуклеосом на смежных участках.

Позиционирование нуклеосом вблизи границы безнуклеосомного участка — процесс стандартный. Если в устройстве нуклеосом допускается вариабельность, например разброс связующих областей по длине на 10 пн, то наведенная специфичность расположения нуклеосом будет убывать по мере удаления от первой нуклеосомы, расположенной у границы. В этом случае позиционирование жестко задано лишь поблизости от границы безнуклеосомного участка ДНК.

Расположение ДНК в нуклеосоме может быть описано двумя способами. В понятие **трансляционного позиционирования** (не имеет никакого отношения к трансляции как синтезу белка) входит описание положения ДНК относительно границ между нуклеосомами; примеры приведены на РИС. 29.33. Трансляционным позиционированием, в частности, определяется, какие последовательности ДНК окажутся в связующих областях и будут более доступными (во всяком случае, если судить по чувствительности к микрококковой нуклеазе).

ДНК центральной частицы нуклеосомы лежит на наружной поверхности гистонового октамера. При этом одна ее сторона погружается в гистоны, в то время как другая сторона молекулы открыта для возможных взаимодействий — например, с регуляторным белком. Поэтому точное положение гистонового октамера по от-

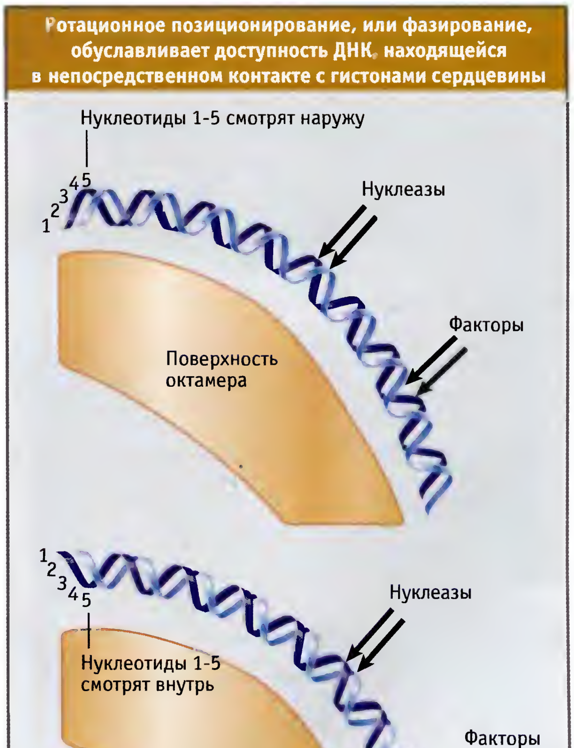


РИС. 29.34 У двойной спирали ДНК, находящейся в контакте с гистонами, ротационное позиционирование определяет сторону, доступную для взаимодействий

ношению к последовательности ДНК может быть очень важным. На РИС. 29.34 изображен эффект **ротационного позиционирования** двойной спирали ДНК относительно поверхности октамера. Если сдвинуть ДНК на неполный виток двойной спирали (<10 пн), паттерн ее доступных сайтов при этом изменится. Чтобы это понять, надо представить себе сдвиг, сопровождаемый вращением ДНК на поверхности октамера.

Трансляционное позиционирование, как и ротационное, может быть важной формой контроля доступа факторов к ДНК. С этой точки зрения, лучше всего описаны случаи специфичного размещения нуклеосом в промоторах. Трансляционное позиционирование и/или специфичное исключение нуклеосом из конкретной последовательности может быть условием, необходимым для формирования транскрипционного комплекса. Связывание с ДНК некоторых регуляторных факторов исключает наличие в этом месте нуклеосомы, и это может привести к образованию границы (точки отсчета) для трансляционного позиционирования нуклеосом. В других случаях регуляторные факторы могут связать ДНК и на поверхности центральной частицы нуклеосомы, но для этого ДНК должна быть так ротационно позиционирована, чтобы ее точки, подходящие для контакта с фактором, находились на стороне двойной спирали, обращенной наружу.

Связь нуклеосомной организации ДНК с транскрипцией мы еще обсудим в разд. 30.4 «Нуклеосомная организация промотора может быть нестандартной», а сейчас лишь отметим, что промоторы (и некоторые другие функционально значимые области геномной ДНК) часто содержат короткие участки, по своей структуре исключающие наличие нуклеосом. Граница такого участка, как правило, служит точкой отсчета позиционирования, и в окрестности этой границы положения нуклеосом относительно последовательности ДНК более или менее фиксированы. Обследование большого участка генома *S. cerevisiae* (картирование 2278 нуклеосом на 482 тпн ДНК) показало, что 60% нуклеосом имеют специфичное расположение в результате граничных эффектов, чаще всего исходящих от промоторов.

29.11 Организованы ли транскрибируемые гены в нуклеосомы?

Ключевые понятия:

- При расщеплении микрококковой нуклеазой транскрибируемых генов нуклеосомы обнаруживаются с той же частотой, что и при расщеплении нетранскрибируемых генов.
- Некоторые интенсивно транскрибируемые гены, в порядке исключения, избавлены от нуклеосомной организации.

Попытки визуализировать гены в процессе их транскрипции привели к противоречивым результатам и непримиримым позициям в их интерпретации, о чем можно судить по следующим двум рисункам.



РИС. 29.35 В препаратах ядрышка транскрипционные единицы генов рибосомальной РНК перемежаются нетранскрибируемыми участками. Воспроизведено по Miller, O. L. and Beatty, B. R. 1969. *Science.* 164: 955–957. Фотография любезна предоставлена Oscar Miller

Интенсивно транскрибируемый хроматин выглядит слишком расправленным, чтобы содержать нуклеосомы. На фото РИС. 29.35 можно видеть интенсивную транскрипцию генов, кодирующих рибосомальную РНК. Разглядеть ДНК мешает крайне высокая плотность посадки РНК-полимераз. Измерить длины транскриптов рРНК напрямую нельзя, поскольку их молекулы упакованы белками, но, зная последовательность рРНК, эти длины можно вычислить. Длина упакованного транскрибируемого участка ДНК, измеряемая как длина «елочки» вдоль оси, составляет ~85% длины рРНК. Это означает, что ДНК почти полностью расправлена.

С другой стороны, транскрипционные комплексы минихромосом SV40, которые можно экстрагировать из инфицированных клеток, содержат стандартный комплект гистонов и выглядят, как бусы. Цепочки РНК можно увидеть торчащими из минихромосомы, примером служит фото РИС. 29.36. Это говорит о том, что транскрипция у SV40 идет, в то время как ДНК пребывает организованной в нуклеосомы. Но, конечно же, ДНК SV40 транскрибируется менее интенсивно, чем гены рибосомальной РНК.

Транскрипция сопровождается расплетанием ДНК и может требовать роспуска всех структур на ограниченных участках хроматина. Транскрипции нужно место, «где развернуться». Характерные облики хромосом типа ламповых щеток и политенных хромосом, описанные в гл. 28 «Хромосомы», умозрительно ассоциируют объемистость структурной организации с экспрессией генов.

Размышляя о транскрипции, следует помнить о соотношении размеров РНК-полимеразы и нуклеосомы. Эукариотические РНК-полимеразы — это большие муль-

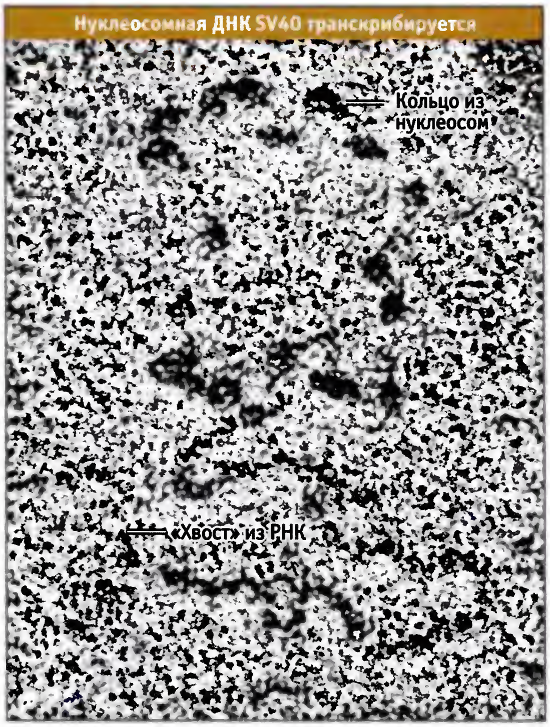


РИС. 29.36 Минихромосома SV40 может транскрибироваться в интактном виде. Воспроизведено по *J. Mol. Bio.*, vol. 131, Gariglio, P., et al., The template of the isolated . . . , p. 131. Авторские права 1979, с разрешения Elsevier. Фотография любезно предоставлена Pierre Chambon

тисубъединичные белки. РНК-полимераза весит, как правило, свыше 500 кДа, в то время как нуклеосома — всего ~260 кДа; наглядное сравнение этих объектов проведено на РИС. 29.37. Даже без детального знания того, как происходит их взаимодействие, ясно, что в нем участвуют два тела, сравнимых по размерам.

Посмотрим на те два оборота, которые делает ДНК вокруг нуклеосомы. Будет ли доступ РНК-полимеразы к ДНК достаточным, если даже при собственной транскрипции нуклеиновая кислота обречена на следование этой траектории? Во время транскрипции, когда РНК-полимераза движется вдоль матрицы, она крепко связывается с участком в ~50 пн, в который входит и локально расплетенный сегмент из ~12 пар оснований. Необходимость расплести ДНК кажется несовместимой с тем, что сегмент, с которым непосредственно работает РНК-полимераза, может сохранять сцепление с поверхностью нуклеосомы. Структурные изменения должны неизбежно сопутствовать транскрипции, и первый вопрос, возникающий по поводу структуры активных генов, таков: остается ли транскрибируемая ДНК организованной в виде нуклеосом? Если в процессе транскрипции гистоновые октамеры смещаются, остаются ли они каким-то образом прикрепленными к транскрибируемой ДНК?

Один из экспериментальных подходов состоит в расщеплении хроматина микрококковой нуклеазой с последующим использованием пробы, специфичной к какому-нибудь конкретному активному гену или генам, чтобы определить, встречаются ли их фрагменты на обычной 200 пн-«лестнице» (см. разд. 29.3 «Геномная ДНК свернута в виде массивов нуклеосом») в ожидаемых концентрациях. Выводы, которые можно сделать из этих экспериментов, довольно ограничены по смыслу, но важны по теоретической значимости. *Транскрибируемые гены по «густоте» содержания нуклеосом эквивалентны последовательностям, которые не транскрибируются.* Таким образом, чтобы быть транскрибированными, генам не обязательно принимать форму организации, альтернативную нуклеосомной. *Со среднестатистическим транскрибируемым геном в каждый момент времени работает одна РНК-полимераза, так, знание общего количества нуклеосом на всем протяжении гена никак не отражает того, что происходит в сайтах, находящихся в непосредственном контакте с ферментом.* Возможно, нуклеосомы там и сохраняются, но более вероятно, что там, где в данный момент времени находится РНК-полимераза, нуклеосома отсутствует, а на участке гена, через который фермент прошел, немедленно восстанавливается нуклеосомная организация.

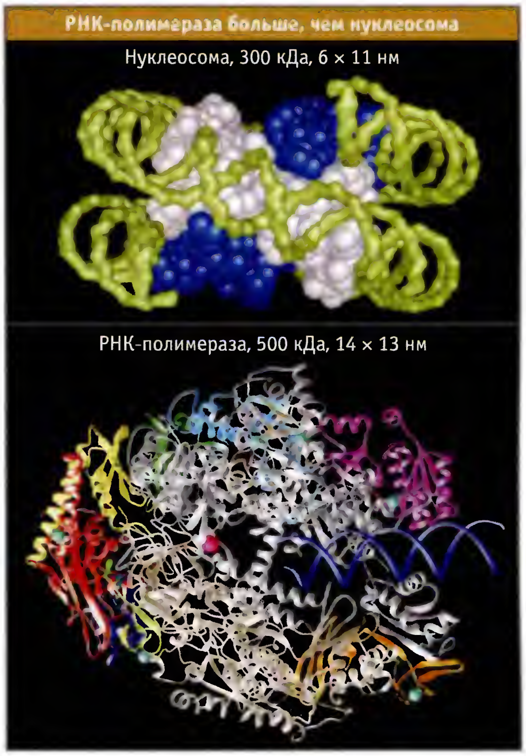


РИС. 29.37 РНК-полимераза по своим размерам сопоставима с нуклеосомой; ей может быть затруднительно следовать за ДНК по поверхности гистонового октамера. Верхний рисунок любезно предоставлен E. N. Moudrianakis, Johns Hopkins University. Нижний рисунок любезно предоставлен Roger Kornberg, Stanford University School of Medicine

29.12 Гистоновые октамеры смещаются транскрипцией

Ключевые понятия:

- В модельной системе РНК-полимераза смещает с ДНК гистоновые октамеры, но, как только фермент пройдет, гистоны и ДНК снова ассоциируют
- Когда через ген проходит транскрипция, его нуклеосомы реорганизуются.

Может ли РНК-полимераза прочесть ДНК прямо на нуклеосоме? Эксперименты показали, что акт транскрипции смещает с ДНК гистоновый октамер. На РИС. 29.38 показано, что происходит, когда РНК-полимераза фага T7 транскрибирует *in vitro* короткий фрагмент ДНК, с которым ассоциирована гистоновая сердцевина одной-единственной нуклеосомы. По завершении транскрипции сердцевина (октамер) по-прежнему ассоциирована с ДНК, но уже в другом месте (как правило, той же самой молекулы).

На РИС. 29.39 изображена модельная схема продвижения РНК-полимеразы. Когда полимераза входит в пределы нуклеосомы, то начинает снимать с нее ДНК, и в какой-то момент достигает точки, в которой ДНК прогибается, снова цепляется за октамер и закрывает фермент в петле. Полимераза, продвигаясь и расплетая ДНК, создает внутри петли положительную сверхспирализацию. Торсионное напряжение, которое при этом возникает в закрытой петле, может быть довольно высоким, поскольку размер петли всего ~80 пн и проход полимеразы через каждую последующую пару оснований сопровождается заметным приращением сверхспирализации. Первые 30 пн в пределах нуклеосомы полимераза проходит легко. Затем ход фермента замедляется, как будто в своем продвижении он сталкивается со все возрастающим сопротивлением. Через каждые 10 пн происходит остановка; вероятно, структура петли требует ротации фермента относительно ДНК на каждом витке двойной спирали. Когда РНК-полимераза достигает середины нуклеосомы (оснований ДНК, лежащих на оси симметрии) она перестает останавливаться и начинает продвигаться быстро. Очевидно, на середине нуклеосомы РНК-полимераза вынуждает октамер сойти с ДНК — возможно, плотность положительных сверхвитков достигает некоторого критического уровня, при котором октамер отторгается. При этом торсионное напряжение двойной спирали впереди полимеразы спадает, что позволяет ферменту свободно продвигаться. Октамер связывается с ДНК позади полимеразы и больше не представляет для нее препятствия. Похоже, что в процессе

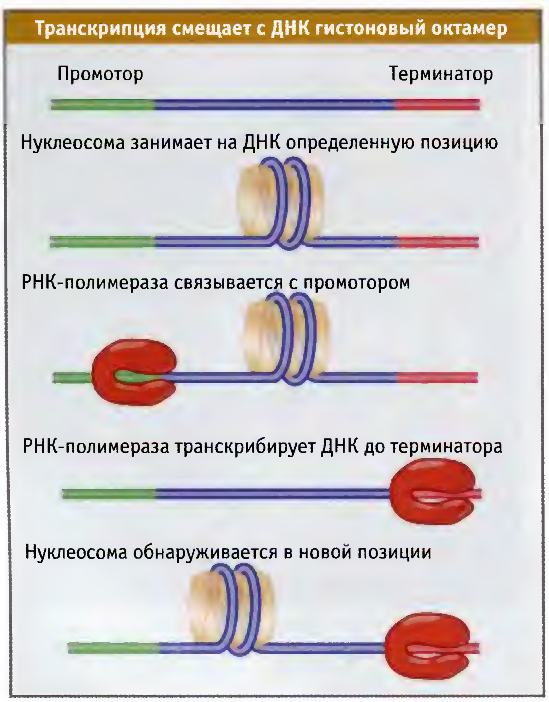


РИС. 29.38 Эксперименты показали, что гистоновый октамер, смещенный с ДНК при транскрипции, может связаться с той же ДНК уже в другой позиции

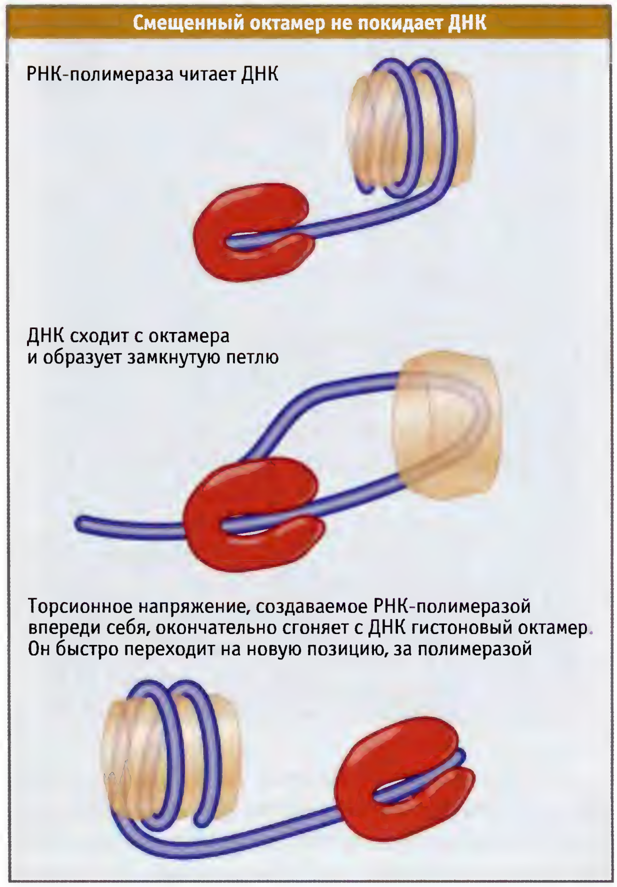


РИС. 29.39 По мере своего продвижения РНК-полимераза снимает ДНК с очередного гистонового октамера. При этом ДНК делает изгиб назад и прикрепляется к полимеразе (или к октамеру) с образованием замкнутой петли. Полимераза продолжает свое движение уже в этой петле и тем самым придает впередилежащей ДНК положительную сверхспирализованность. Не выдержав, октамер, без отрыва от ДНК и/или полимеразы, переходит на новую позицию, уже прочтенную полимеразой

смены своего положения на ДНК октамер непрерывно сохраняет с ней какую-то форму контакта.

Высвобождается ли октамер как интактная единица, или его целостность при этом страдает? Искусственные сшивки белков октамера транскрипции не препятствуют. Транскрипция может продолжаться, даже если сшивками охвачены центральные домены гистонов сердцевины. Создается впечатление, что для транскрипции не требуется ни диссоциации, ни даже сколько-нибудь капитального демонтажа октамера. Однако добавление гистона H1 быстро приводит систему к упадку транскрипционной активности. Напрашивается два вывода: гистоновый октамер (присутствует ли он или смещен) функционирует как целостная единица; октамер может быть необходим для устранения H1 из активного хроматина или может выступать в роли модулятора взаимодействий этого гистона с активным хроматином.

Итак, сравнительно небольшая вирусная РНК-полимераза может во время транскрипции сместить одну нуклеосому, которая позади фермента формируется заново. В эукариотическом ядре ситуация, конечно же, гораздо сложнее. РНК-полимераза там гораздо крупнее (чем у фага T7), а на пути ее стоит целая очередь связанных нуклеосом. Преодоление этого препятствия требует участия дополнительных факторов, воздействующих на хроматин (см. гл. 30 «Контроль структуры хроматина»).

Организация нуклеосом может быть изменена транскрипцией. На РИС. 29.40 показано, что происходит с дрожжевым геном *URA3*, когда он транскрибируется под контролем индуцируемого промотора. Позиционирование нуклеосом проверяется при помощи микрококковой нуклеазы, путем выявления расположения доступных для нее сайтов относительно фиксированной точки — сайта рестрикции в 5′-концевой области гена. Исходно в гене выявляется паттерн нуклеосом, охватывающий значительное расстояние: от промотора до почти 3′-концевой области (в которой позиционирование сходит на нет). После того как ген начал экспрессироваться, на смену паттерну полос, отображающему позиционирование нуклеосом, возникает шмер (по суммарной интенсивности — такой же). Это означает, что плотность нуклеосом остается прежней, но теперь, у транскрибируемых копий гена, они расположены в разных фазах. Транскрипция уничтожает позиционирование нуклеосом. Через 10 мин после репрессии позиционирование восстанавливается (хоть и не полностью). Эти результаты также означают, что позиции нуклеосом могут быть скорректированы без репликации.

В общем случае, обоснованным является предположение, что по мере своего продвижения РНК-полимераза смещает гистоновые октамеры. Если ДНК позади полимеразы доступна, октамер закрепляется там. Скорее всего, октамер никогда не теряет всех своих контактов с ДНК. Как ему это удается — загадка, потому что он нисколько не демонтируется, в то время как мимо него вдоль ДНК движется объект (РНК-полимераза) еще более крупный, чем сам октамер. Может быть, октамер «проталкивается назад» посредством установления контактов с РНК-полимеразой. Если ДНК позади полимеразы не доступна (например, из-за того, что вслед за первой полимеразой идет вторая), то октамер может быть перманентно устранен, и ДНК останется в расправленной форме.

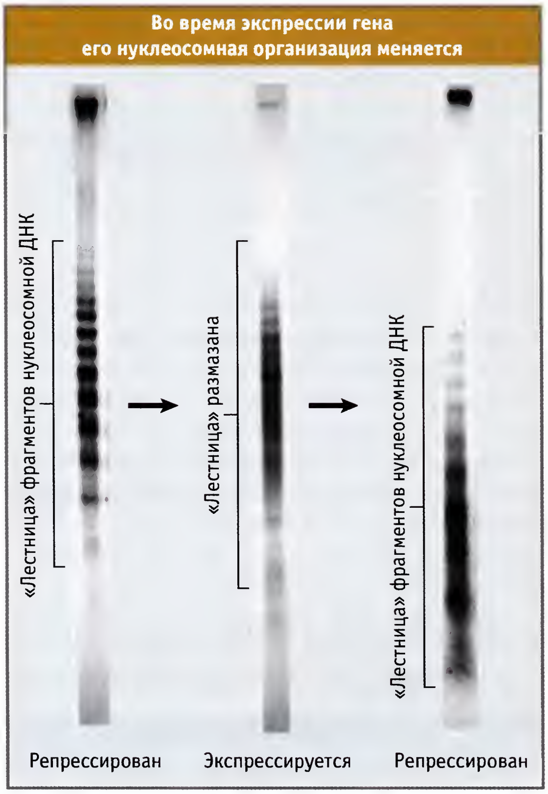


РИС. 29.40 В данном эксперименте последовательность гена *URA3*, продолженная последовательностью рибосомальной ДНК, была поставлена под регулируемый промотор *GAL1*. Перед транскрипцией последовательность *URA3* организована в виде позиционированных нуклеосом. После индукции транскрипции позиции нуклеосом приобретают случайный характер. Если репрессировать транскрипцию, позиционирование восстановится. Воспроизведено по Suter, B., et al. 1997. EMBO J. 16: 2150–2160. Авторские права © Oxford University Press. Фотография любезно предоставлена Fritz Thomas, ETH Zürich

29.13 Смещение нуклеосом и реассоциация требуют участия специальных факторов

Ключевые понятия:

- Вспомогательные факторы требуются как РНК-полимеразе, чтобы смещать гистоновые октамеры во время транскрипции, так и гистонам, чтобы снова формировать нуклеосомы после транскрипции.

Смещение нуклеосом с ДНК — это ключевое требование для всех этапов транскрипции. Лучше всего процесс смещения описан для этапа инициации. Активные промоторы отмечены сайтами, гиперчувствительными к ДНКазе, поскольку в этих сайтах гистоновые октамеры оказались смещены с ДНК (см. разд. 29.18 «Сайты, гиперчувствительные к ДНКазе, реагируют на изменения в структуре хроматина»). Для устранения октамеров требуются ремоделирующие комплексы. Эти белковые комплексы используют энергию гидролиза АТФ, чтобы изменять структуру хроматина; в промоторы их рекрутируют факторы транскрипции (см. разд. 30.4 «Нуклеосомная организация промотора может быть нестадартной»). Это означает, что РНК-полимераза начинает синтез РНК на коротком отрезке ДНК, свободном от нуклеосом. Чтобы полимераза продолжила свой ход на этапе элонгации, нужно убирать с ее пути гистоновые октамеры. Во избежание появления оголенной ДНК, после транскрипции нуклеосомы должны формироваться заново.

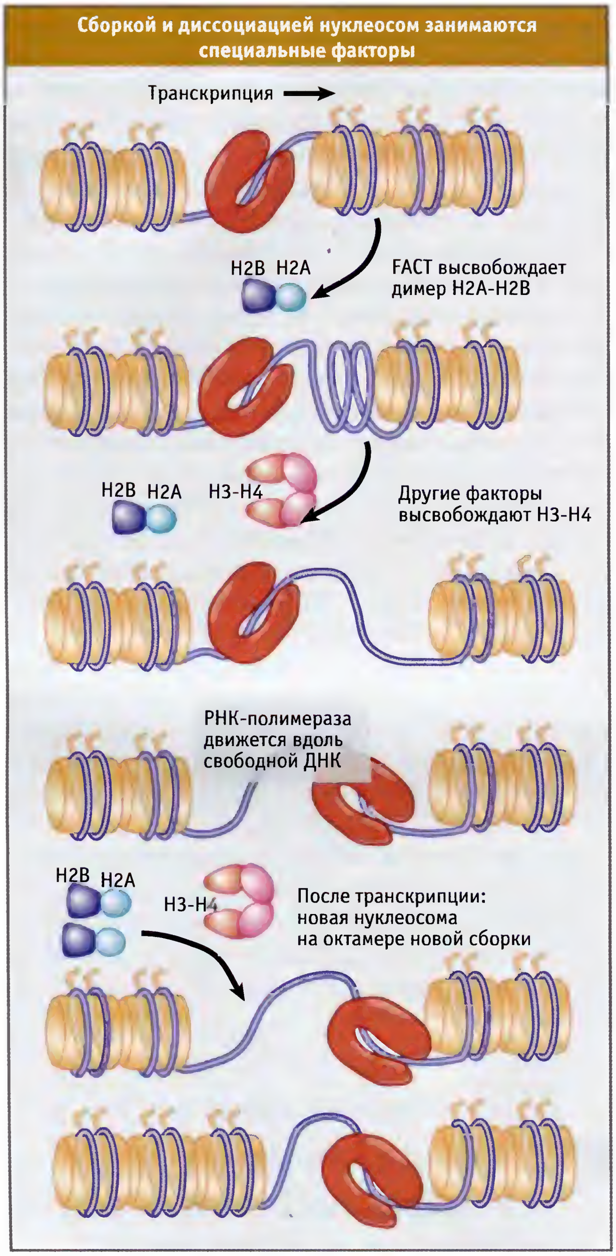


РИС. 29.41 Чтобы нуклеосомы не мешали транскрипции, впереди полимеразы полным ходом идет разборка гистоновых октамеров. Скорее всего, разборка начинается с ухода димеров H2A-H2B. Для восстановления нуклеосом после транскрипции октамеры собираются заново

Для транскрипции РНК-полимеразой II *in vitro* требуется белок, называемый FACT (facilitates chromatin transcription; «способствует транскрипции хроматина»), который ведет себя как фактор элонгации транскрипции. Не являясь частью РНК-полимеразы, он специфично ассоциирует с ней на время элонгации. FACT состоит из двух очень консервативных субъединиц и ассоциирован с хроматином активных генов.

Если добавить FACT к изолированным нуклеосомам, он заставит их потерять димеры H2A–H2B. Во время транскрипции *in vitro* FACT переводит нуклеосомы в «гексасомы», потерявшие по одному димеру H2A–H2B. Это указывает на FACT, как на элемент механизма смещения октамеров во время транскрипции. Также FACT может участвовать в восстановлении нуклеосом после транскрипции, поскольку он содействует их формированию *in vitro* (из ДНК и гистонов сердцевины).

На основании этих результатов можно построить модель, изображенную на РИС. 29.41. Впереди РНК-полимеразы FACT открепляет от нуклеосомы димер H2A-H2B, а затем помогает добавить его к нуклеосоме, восстанавливающейся позади полимеразы. Чтобы довести этот процесс до конца, требуются, должно быть, и другие факторы. FACT же участвует и в других событиях, ход которых требует смещения нуклеосом, в том числе — репликации и репарации ДНК.

Для поддержания целостности хроматина на участках, где идет транскрипция, требуются и другие факторы, и, возможно, они тоже имеют отношение к разборке и сборке нуклеосом, но детальной информации об этих факторах пока нет.

29.14 Инсуляторы блокируют действие энхансеров и гетерохроматина

Ключевые понятия:

- Инсуляторы способны заблокировать любые активирующие и инактивирующие эффекты, порождаемые энхансерами, сайленсерами или LCR.
- Инсуляторы могут служить барьерами, пресекающими распространение гетерохроматина.

Элементы геномной ДНК, препятствующие прохождению (через себя и вдоль ДНК) активирующих либо инактивирующих сигналов, называют **инсуляторами**. Они обладают как минимум одним из двух ключевых свойств:

- Если инсулятор расположен между энхансером и промотором, *он предотвращает активацию промотора энхансером.* Блокирующий эффект изображен



РИС. 29.42 Энхансер активирует промотор в своей окрестности; действие энхансера может заблокировать инсулятор, расположенный между ним и промотором



РИС. 29.43 Распространяясь по хромосоме, гетерохроматин может захватывать и блокировать активные промоторы. Инсулятор, который служит барьером для распространения гетерохроматина, позволяет промотору оставаться активным

на РИС. 29.42. Из рисунка видно, почему действие энхансера ограничивается конкретным промотором.

- Если инсулятор расположен между активным геном и гетерохроматином, *он служит барьером, защищающим ген от инактивирующего эффекта, исходящего от гетерохроматина.* (Гетерохроматин — это участок хроматина, неактивный из-за высокой компактизации; см. разд. 31.2 «Гетерохроматин распространяется от метса зарождения».) Эффект барьера изображен на РИС. 29.43.

Некоторые инсуляторы обладают обоими функциональными свойствами, блокирующим и барьерным, другие — только одним из них (либо обоими, но в режиме разделения во времени). Обе функции опосредованы изменениями структуры хроматина, но эти изменения могут быть различными. Все же в любом конкретном случае инсулятор служит пределом распространения дальнодействующих эффектов.

Зачем нужны инсуляторы? Их главное предназначение состоит в противодействии неизбирательному влиянию энхансеров на промоторы. Большинство энхансеров активно в отношении даже очень отдаленного промотора. Инсулятор может сузить действие энхансера, четко ограничив район распространения его эффекта (где энхансер будет специфично действовать на конкретный промотор). Похожим образом инсулятор может защитить ген, расположенный вблизи гетерохроматина, от нечаянной инактивации (в случае если гетерохроматин начнет распространяться). Инсуляторы, таким образом, функционируют в качестве элементов генома, повышающих точность регуляции экспрессии генов.

29.15 Инсуляторы могут ограничивать домен

Ключевые понятия:

- **В инсуляторах, из-за особой структуры хроматина, есть сайты, гиперчувствительные к ДНКазе. Два инсулятора могут защитить участок, расположенный между ними, от влияний геномного окружения.**

Инсуляторы были открыты при анализе геномной области *Drosophila melanogaster*, изображенной на РИС. 29.44. Два гена, кодирующих белок Hsp70 (белок теплового шока 70) лежат внутри участка в 18 тпн, соответствующего полосе 87A7 политенного набора хромосом. По краям полосы хроматин имеет особую структуру; там находятся последовательности *scs* и *scs'* (specialized chromatin structures; «специализированные структуры хроматина»). Каждый из элементов, *scs* и *scs'*, представляет собой участок высокой устойчивости к разрушению ДНКазой I, фланкированный несколькими сайтами, гиперчувствительными к ДНКазе и разнесенными на ~100 пн. При активации генов Hsp70 тепловым шоком меняется паттерн расщепления ДНК по этим сайтам.

Элементы *scs* и *scs'* «инсулируют» (изолируют) гены *hsp70* от влияния их геномного окружения. Если взять элементы *scs* и поместить по обе стороны гена *white*, тот сможет функционировать в любом месте генома, где бы

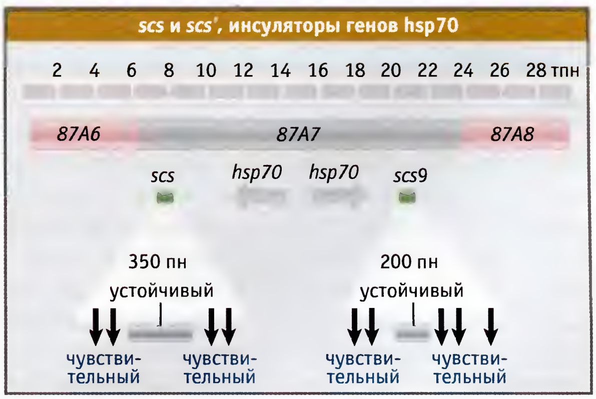


РИС. 29.44 В геноме *D. melanogaster* есть домены, ограниченные элементами *scs* и *scs'*. Края этих доменов характеризуются особой структурой хроматина с гиперчувствительными сайтами ДНК. Элементы *scs* и *scs'* оберегают гены, расположенные между ними, от эффектов геномного окружения



РИС. 29.45 В политенных хромосомах *Drosophila* белок BEAF, специфично связывающий инсулятор *scs'*, в большом количестве находится в промежутках между полосами. В верхней и нижней частях изображения красный цвет обозначает ДНК (полосы); в верхней части изображения зеленый цвет обозначает белок BEAF32, которого много в промежутках. Желтый цвет появляется там, где зеленая и красная метки совпадают, то есть там, где белок BEAF32 находится в полосе. Воспроизведено по *Cell*, vol. 81, Zhao, K., Hart, C. M., and Laemmli, U. K., Visualization of chromosomal . . . , pp. 879–889. Авторские права 1995, с разрешения Elsevier. Фотография любезно предоставлена Ulrich K. Laemmli, University of Geneva, Switzerland

он ни очутился (даже там, где он, по идее, должен быть репрессирован геномным контекстом — например, в гетерохроматине).

Элементы *scs* и *scs'*, по-видимому, не оказывают на экспрессию генов самостоятельного контролирующего действия, ни положительного, ни отрицательного, а просто препятствуют сообщению между участками генома, ограничивая распространение контролирующих эффектов вдоль ДНК. Если ген граничит с участками, которые его репрессируют, элементы *scs* могут понадобиться, чтобы заблокировать приход ингибирующих эффектов с флангов (участков, примыкающих к гену) и могут, таким образом, быть принципиально значимыми для его экспрессии. Делеция элемента *scs* может полностью выключить экспрессию генов на соседних участках.

У *scs* и *scs'* разная первичная структура, и принцип изолирующей активности у этих элементов, по-видимому, различается. Ключевая последовательность элемента *scs* состоит из 24 пн и связывает продукт гена *zw5*. Сущность *scs'*, как инсулятора, заключена в серии повторов CGATA. Повторы связывают группу родственных белков, называемых BEAF-32. Распределение этих белков в ядре дискретно, и самое замечательное — это их локализация в политенных хромосомах. На РИС. 29.45 показано, что антитела к BEAF-32 окрашивают ~50% промежутков между полосами политенных хромосом. Это означает, что в геноме много инсуляторов, и что BEAF-32 — стандартный компонент их аппарата. Полосы при этом вырисовываются в качестве функциональных подразделений политенной хромосомы, а промежутки между полосами часто несут в себе инсуляторы, обеспечивающие функциональную непроницаемость.

Другой пример инсулятора, ограничивающего хромосоный домен, находится в β-глобиновом LCR цыпленка. Этот LCR (опознаваемый как группа сайтов, гиперчувствительных к ДНКазе) контролирует экспрессию всех β-глобиновых генов (см. разд. 29.20 «LCR может контролировать домен»). Самый левый гиперчувствительный сайт β-глобинового LCR, HS4, является инсулятором; он маркирует 5'-конец функционального домена и ограничивает зону действия LCR глобиновыми генами домена.

Ген, окруженный инсуляторами, обычно защищен от вторжения инактивирующих эффектов со стороны геномного контекста. Чтобы доказать это утверждение, можно выполнить опыт с трансфекцией ДНК в клеточные геномы по случайным местоположениям. Режим экспрессии гена в таких вставках может быть любым: где-то в геноме он будет экспрессироваться активно, где-то — не экспрессироваться вовсе. Но если края вставки снабдить инсуляторами со свойствами барьеров, то экспрессия будет идти с одной и той же равномерной интенсивностью, вне зависимости от того, в какой сайт генома попадет эта вставка.

29.16 Инсулятор транспозона *gypsy* и его белковый аппарат

Ключевые понятия:

- Некоторые инсуляторы обладают избирательностью — они могут останавливать распространение эффекта только в одном направлении.

Рассмотрим вставки транспозона *gypsy* в локус *yellow* (*y*) *Drosophila melanogaster*. В разных тканях локус *y* регулируется одним из четырех энхансеров, как показано на РИС. 29.46, и вставка *gypsy* приводит к потере функции этого гена в некоторых тканях. Внедренный *gypsy* блокирует действие всех энхансеров, отделяемых им от промотора. Этот эффект обусловлен последовательностью инсулятора, расположенной в одной из концевых областей транспозона. Инсулятор работает вне зависимости от того, как ориентирована вставка.

Белки, работающие на инсуляторе *gypsy* были идентифицированы как продукты двух других локусов, от которых работа инсулятора зависит в *транс*-манере (то

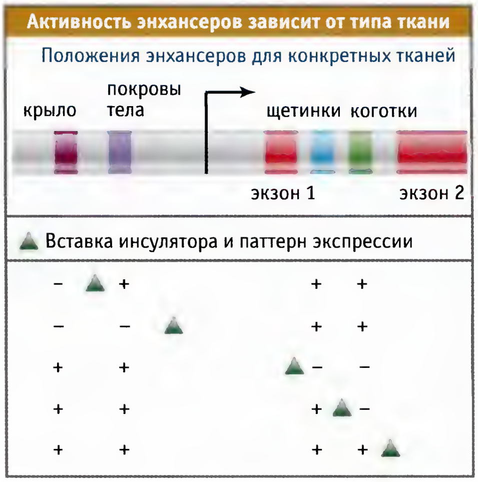


РИС. 29.46 **Инсулятор транспозона *gypsy* блокирует действие энхансера, если его поместить между энхансером и промотором**

есть их местоположения в геноме не существенны). Мутации по локусу *su(Hw)* гасят функцию инсулятора *gypsy*. Это обстоятельство послужило ключом к установлению взаимодействия между белком, кодируемым *su(Hw)*, и *gypsy*. Белок узнает последовательность инсулятора при помощи аминокислотного мотива цинковый палец; его иммуногистохимическое картирование на политенных хромосомах показало, что он связывает большое число сайтов. Инсулятор *gypsy* содержит двенадцать копий последовательности в 26 пн, которую и связывает su(Hw).

Второй важный локус называется *mod(mdg4)*; мутации по нему производят противоположный эффект. Эти мутации принципиально расширяют возможности инсулятора: он, строго говоря, перестает быть инсулятором и теперь блокирует действие на промотор энхансеров по обе стороны от себя. По отношению к гену *mod(mdg4)* ген *su(Hw)* является эпистатическим; это означает, что двойные мутанты имеют фенотип *su(Hw)*$^{-}$, и говорит о том, что белок mod(mdg4) действует через белок **su(Hw)**.



РИС. 29.47 **Комплексы Su(Hw)/mod(mdg4) находятся на периферии ядра. Они могут организовывать ДНК в виде петель и тем самым ограничивать взаимодействие энхансеров с промоторами**

Связывание su(Hw) с ДНК, за которым следует связывание mod(mdg4) с su(Hw), блокирует активацию промотора. Инсулятор *gypsy*, связанный с su(Hw), распространяет ингибирующее действие в обе стороны, но по одному из направлений это распространение останавливает mod(mdg4).

Скорее всего, действие любого инсулятора обусловлено внесением изменений в структуру хроматина. Хотя в геноме *Drosophila* сайты связывания с белками su(Hw) и mod(mdg4) встречаются в более чем пяти сотнях местоположений, визуализация этих сайтов показывает, что они ко-локализуются в ~25 дискретных сайтах на периферии клеточного ядра. Это обстоятельство подкрепляет модель работы *gypsy*, изображенную на РИС. 29.47, где молекулы su(Hw), связавшие ДНК в разрозненных сайтах, сходятся вместе благодаря связыванию с mod(mdg4). Комплекс su(Hw)/mod(mdg4) локализован на периферии ядра. Связанная с ним ДНК (вернее, хроматин) имеет вид петель. Среднестатистический комплекс может удерживать на себе ~20 таких петель. Взаимодействие энхансер-промотор работает внутри петли, но не между петлями.

29.17 Инсуляторы бывают разными по силе

Ключевые понятия:

- Инсуляторы могут различаться по эффективности, с которой они блокируют прохождение активирующего сигнала.

На некоторых участках генома элементы ДНК, действующие в *цис*-манере (то есть на мишени, расположенные вдоль той же самой молекулы ДНК), но различные по смыслу своей активности, скомбинированы и образуют регуляторные единицы сложного профиля. Один из таких участков в геноме *Drosophila*, *Fab*-7, маркируется делециями по локусу *bithorax*. Этот локус содержит серию *цис*-действующих регуляторных элементов, контролирующих активность трех единиц транскрипции. Соответствующая часть локуса *bithorax* изображена на РИС. 29.48. Регуляторные элементы *iab-6* и *iab-7* контролируют экспрессию соседнего с ними гена *Abd-B* в двух идущих подряд сегментах (отделах тела) зародыша, А6 и А7. Делеция *Fab*-7 заставляет А6 развиваться по типу А7. Этот эффект доминантен, то есть, в случае делеции *Fab*-7, контроль над *Abd-B* в сегменте А6 переходит от *iab-6* к *iab-7*. В молекулярных терминах это можно интерпретировать так, что *Fab*-7 играет роль границы, препятствующей действию *iab*-7, когда активен элемент *iab*-6.

Подобно другим разграничивающим элементам (инсуляторам), *Fab*-7 структурирован в хроматине особым образом и потому содержит сайты, гиперчувствительные к ДНКазе. Функционально, *Fab*-7 можно подразделить на более мелкие элементы двух типов (анализ проводился

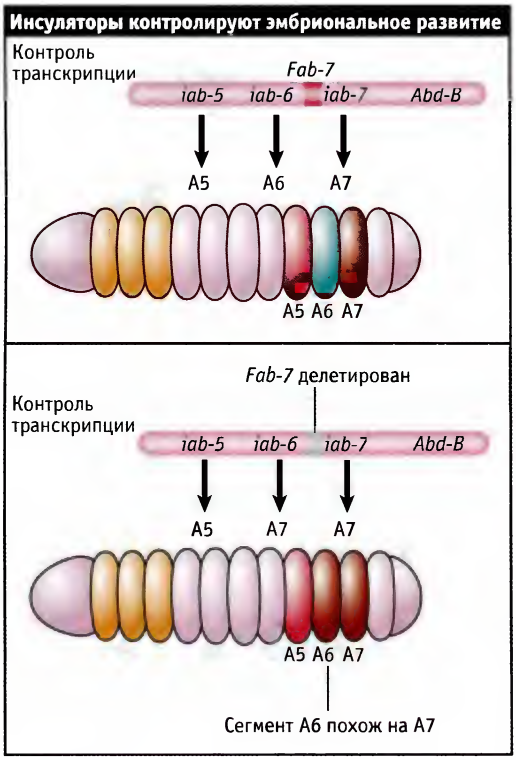


РИС. 29.48 Пограничный элемент *Fab-7* необходим для того, чтобы регуляторные элементы *iab-6* и *iab-7* работали независимо

по результатам внесения в *Fab*-7 небольших делеций и/или тестирования фрагментов на способность служить инсулятором). Последовательность в ~3,3 тпн ведет себя как инсулятор. Последовательность в ~0,8 тпн ведет себя как репрессор, действующий на *iab*-7. Сочетанием этих двух элементов объясняется сложное генетическое поведение *Fab*-7 (которое здесь подробно не описано).

Очень интересными оказались результаты экспериментов, в которых *Fab*-7 заменяли другими инсуляторами. Они заставляют пересмотреть тезис о том, что *Fab*-7 должен просто предотвращать взаимодействие *iab*-6 и *iab*-7. Если заменить *Fab*-7 на последовательность сайта связывания su(Hw), то можно получить дополнительный эффект: *iab*-5 выбивается из-под негативного контроля со стороны *iab*-7. А если вместо *Fab*-7 поставить элемент *scs*, тот же эффект достигнет региона *iab*-4.

Результаты заставляют предположить существование функциональной схемы, согласно которой более активные инсуляторы могут блокировать действие регуляторных элементов на более далеких расстояниях. Такой вывод затрудняет понимание принципов работы разграничивающих элементов. Теперь их действие нельзя объяснить простым блоком передачи регуляторных эффектов через границу на ДНК. Тем не менее это расширило представление о сложнейшей многоуровневой конкурентной системе контроля активности хроматина. В такой системе влиятельность регуляторного элемента ДНК — собственно, белкового аппарата, с которым этот элемент взаимодействует в силу специфики своей нуклеотидной последовательности — может быть эквивалентна дальности его действия, измеренной в парах оснований той же ДНК.

Еще больше усложняет ситуацию существование таких элементов, как **анти-инсуляторы**, которые позволяют энхансеру преодолевать действие инсулятора, стоящего между ним и промотором. Регуляторные взаимодействия участков ДНК предстают в свете взаимодействия с в локальной структурой хроматина, где над каждым уровнем контроля можно выстроить следующий, и т.д.

29.18 Сайты, гиперчувствительные к ДНКазе, реагируют на изменения в структуре хроматина

Ключевые понятия:

- Гиперчувствительные сайты обнаруживаются в промоторах экспрессируемых генов.
- Эти сайты возникают из-за связывания с ДНК факторов транскрипции, смещающих с нее гистоновые октамеры.

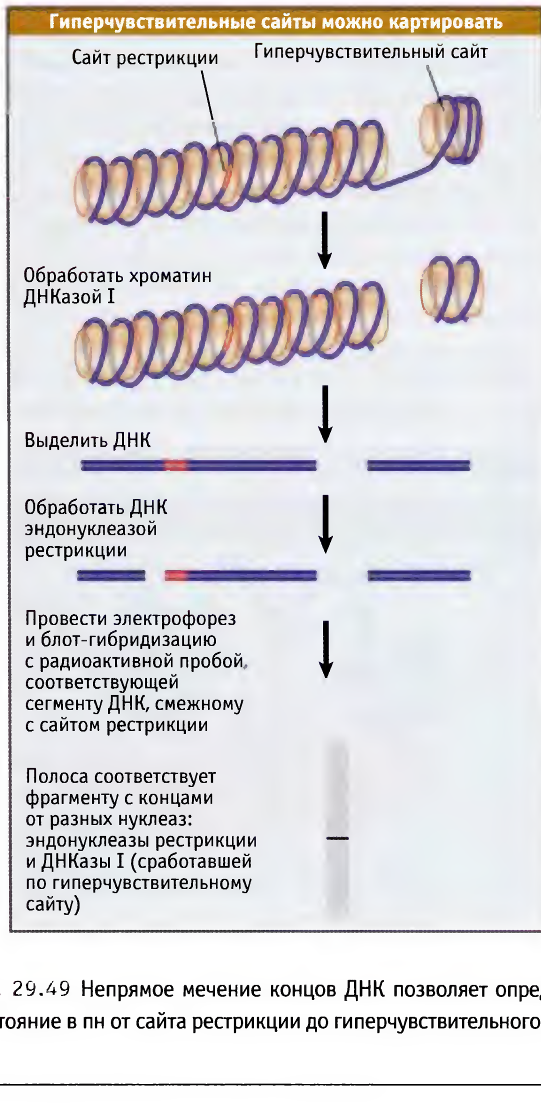


РИС. 29.49 Непрямое мечение концов ДНК позволяет определить расстояние в пн от сайта рестрикции до гиперчувствительного сайта



РИС. 29.50 Мини-хромосома SV40 содержит область, где нет нуклеосом. Фотография любезно предоставлена Moshe Yaniv, Pasteur Institute

В дополнение к широкомасштабным изменениям, на активных (или потенциально активных) участках хроматина происходят локальные изменения его структуры. Они имеют место в специфических сайтах, причастных к инициации транскрипции, или там, где ДНК обладает определенными особенностями первичной структуры. Существование такого рода изменений было обнаружено при отслеживании процесса расщепления хроматина ферментом ДНКазой I в очень низкой концентрации.

При расщеплении хроматина ДНКазой I, разрывы вносятся в ДНК первым делом по определенным **гиперчувствительным сайтам**. Принято считать, что в этих сайтах ДНК особенно уязвима оттого, что не организована, как обычно, в виде нуклеосом. Типичный гиперчувствительный сайт в ~100 раз более чувствителен к атаке фермента, чем основная масса хроматина. Положения сайтов могут быть определены при помощи методики непрямого концевого мечения, с которой мы недавно познакомились при обсуждении позиционирования нуклеосом.

Применительно к данному случаю методика проиллюстрирована на РИС. 29.49. Одному из концов детектируемого фрагмента соответствует двухцепочечный разрыв по гиперчувствительному сайту, вносимый ДНКазой I. Другой конец фрагмента порождает рестрикция, сайт для которой занимает фиксированное положение в последовательности геномной ДНК.

Многие гиперчувствительные сайты имеют отношение к экспрессии генов. Любой активный ген имеет, по меньшей мере, один такой сайт в промоторе. *Как правило, гиперчувствительные сайты возникают только в тех клетках, где «их» ген, с которым они ассоциированы, экспрессируется*; когда ген не активен, те же самые сайты не проявляют никакой гиперчувствительности к ДНКазе. 5′-концевые гиперчувствительные сайты возникают еще до начала транскрипции, а последовательности этих сайтов — по данным мутационного анализа — для экспрессии *необходимы*.

Гиперчувствительные сайты минихромосомы вируса SV40 изучены особенно хорошо. На сайт возле ориджина репликации, чуть выше промотора поздней единицы транскрипции, предпочтительно действуют ДНКаза I, микрококовая нуклеаза, другие нуклеазы (в том числе ферменты рестрикции).

Состояние минихромосомы SV40 можно визуализировать средствами электронной микроскопии. В значительной доле образцов (до 20%) можно разглядеть «пробел» в нуклеосомной организации материала; примером служит фото РИС. 29.50. Этот «пробел» — участок длиной ~120 нм, или около 350 пн — с обеих сторон окружен нуклеосомами. Видимая зона «пробела» содержит сайт, гиперчувствительный к нуклеазам. Это прямое свидетельство взаимосвязи гиперчувствительности с отсутствием нуклеосом.

Гиперчувствительный сайт не обязательно является таковым по отношению ко всем нуклеазам в равной степени. Карты двух гиперчувствительных сайтов приведены на РИС. 29.51.

Гиперчувствительный сайт внутри безнуклеосомного «пробела» на минихромосоме SV40, на самом деле, поделен надвое: внутри него есть участок, защищенный от нуклеаз, предположительно, благодаря тому, что связан с негистоновым белком (белками). Некоторые элементы ДНК «пробела» *необходимы* для работы промотора.

Гиперчувствительному сайту, расположенному в промоторе β-глобиновых генов цыпленка, оказывают предпочтение несколько нуклеаз: его, куда «охотнее» других участков ДНК, разрезают ДНКаза I, ДНКаза II и микрококковая нуклеаза. Производя расщепление, эти ферменты выбирают разные точки внутри сайта. Когда ген (потенциально) активен, нуклеазы (в совокупности) имеют доступ к участку β-глобиновой ДНК с координатами от −270 до −70.

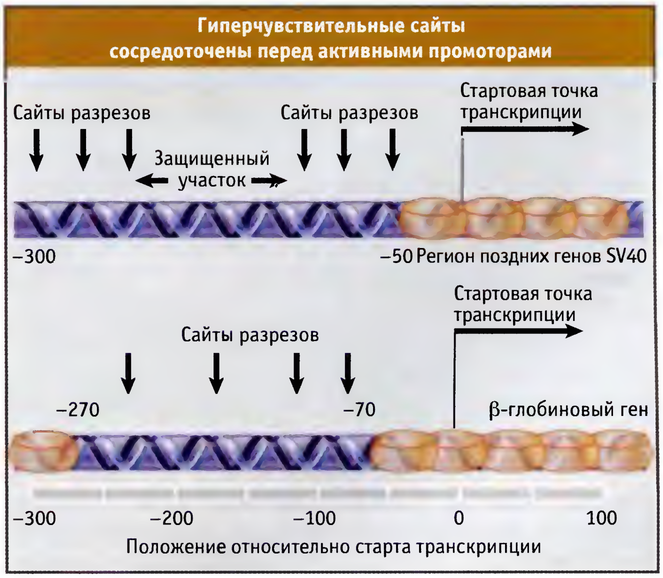


РИС. 29.51 В безнуклеосомной области SV40 есть гиперчувствительные сайты, участки повышенной чувствительности, а также защищенный участок ДНК. Гиперчувствительный сайт β-глобинового гена курицы находится на участке, чувствительном сразу к нескольким нуклеазам

Как устроен гиперчувствительный сайт? Его первоочередная доступность для нуклеаз говорит об отсутствии на нем гистоновых октамеров (нуклеосом). Это не означает, что сайт вообще «чист»: без помощи негистоновых белков исключить из ДНК нуклеосомы, чаще всего, невозможно (не говоря уже о том, что оголенный участок ДНК всегда находится под угрозой повреждения). Гиперчувствительный сайт возникает в результате специфичного связывания с ДНК регуляторных белков, вытесняющих оттуда нуклеосомы. При посадке на ДНК такой белок закрывает ее собой; именно поэтому какая-нибудь часть гиперчувствительного сайта, как правило, защищена от нуклеаз.

Белки, присутствием которых на ДНК вызвано наличие в ней гиперчувствительного сайта, — это разного рода регуляторные факторы. Гиперчувствительные сайты возникают не беспорядочно, а в промоторах, ориджинах, центромерах и на других функционально значимых участках генома. Иногда их наличие сопровождается развитой многоуровневой структурированностью хроматина. Граница гиперчувствительного сайта может служить точкой отсчета для позиционирования серии нуклеосом. Возникновение гиперчувствительного сайта в связи с транскрипцией может быть обусловлено фактором транскрипции, который связывается с промотором, подготавливая его к посадке РНК-полимеразы (см. разд. 30.4 «Нуклеосомная организация промотора может быть нестандартной»).

Насколько стабильными бывают гиперчувствительные сайты, можно оценить по их поведению в фибробластах цыпленка, трансформированных термочувствительным опухолевым вирусом. Эта система обладает тем необычным свойством, что, хотя, по типу своей дифференцировки, фибробласты не принадлежат к эритроидному ряду, их трансформация указанным вирусом при нормальной температуре приводит к активации глобиновых генов. (Активированные глобиновые гены, как известно, имеют гиперчувствительные сайты.) Если провести трансформацию при более высокой (не пермиссивной) температуре, то глобиновые гены не активируются, и гиперчувствительные сайты в них не возникают. Наконец, если глобиновые гены были активированы трансформацией при низкой температуре, их можно инактивировать, подняв температуру. При этом гиперчувствительные сайты в них будут сохраняться на протяжении, по меньшей мере, двадцати последующих клеточных циклов.

Этот результат говорит о том, что обретение гиперчувствительного сайта — это лишь один из индикаторов изменений, которые должны произойти с геном, чтобы его транскрипция инициировалась. Другое следствие состоит в том, что возникновение гиперчувствительного сайта и его последующее поддержание имеют не одну и ту же природу. В данной системе, если появился гиперчувствительный сайт, он будет поддерживаться и воспроизводиться при репликации, несмотря на отсутствие транскрипции (и условий, необходимых для ее индукции). Возможно ли удалить эти сайты каким-нибудь специфическим вмешательством?

29.19 Участки хроматина, содержащие активные гены, обладают повышенной чувствительностью к нуклеазам

Ключевые понятия:

- **Домен, в котором есть транскрибируемый ген, отличается повышенной чувствительностью к ДНКазе I.**

Области генома, содержащей активный ген, присуща измененная структура. Изменение структуры предшествует транскрипции; это совсем не то же самое, что перераспределение нуклеосом в связи с продвижением РНК-полимеразы вдоль ДНК.

Один из признаков изменения структуры транскрибируемого хроматина — это его повышенная чувствительность к ДНКазе. На основании повышенной чувс-



РИС. 29.52 Чувствительность к ДНКазе I можно измерить как скорость исчезновения материала, гибридизующегося с конкретной пробой

твительности к ДНКазе в общей массе хроматина можно выделить хромосомный **домен**, то есть область с измененной структурой, где есть, по меньшей мере, одна активная единица транскрипции. Обратите внимание, термин «хромосомный домен», в контексте проблематики этого раздела, не имеет прямого отношения к петлеобразной структуре в составе хроматина или хромосомы, обозначавшейся этим термином ранее.

Когда хроматин расщепляют ДНКазой, его ДНК постепенно деградирует до материала, растворимого в кислой среде (то есть до очень коротких фрагментов). Общий ход реакции можно отслеживать как процент ДНК, перешедший в раствор. *Если в основной своей массе хроматин переходит в раствор лишь на 10%, то для хроматина по местонахождению активного гена эта доля составляет более 50%.* Активные гены, таким образом, деградируют быстрее.

Об участи конкретных генов, предоставленных действию ДНКазы, можно судить по измерению количества их ДНК, просуществовавший до гибридизации со специфичной пробой. Протокол методики и типичные результаты ее применения показаны на РИС. 29.52. Исчезновение на «картинке» конкретной полосы после щадящей обработки хроматина ДНКазой говорит о том, что соответствующий участок деградировал и, следовательно, обладал повышенной чувствительностью к ДНКазе.

На РИС. 29.53 показано, что происходит с β-глобиновыми генами и с геном овальбумина в хроматине, экстрагированном из эритроцитов цыпленка, при обработке ДНКазой. (В птичьих эритроцитах, красных кровяных тельцах, глобиновые гены экспрессируются, а ген овальбумина — нет.) Фрагменты β-глобиновых генов быстро измельчаются ДНКазой и теряются за то время, пока фрагменты гена овальбумина успевают лишь слегка деградировать. На самом деле, ген овальбумина расщепляется ДНКазой со скоростью, типичной для ДНК основной массы хроматина.

Таким образом, хроматин, в своей массе, довольно-таки устойчив к расщеплению ДНКазой I, поскольку содержит в основном неактивные последовательности, в том числе, нетранскрибируемые гены. *Специфически уязвимым для ДНКазы ген становится в той ткани (тканях), где идет его экспрессия.*

Является ли такое свойство, как приоритетное расщепление нуклеазой, характерным лишь для генов, экспрессируемых достаточно интенсивно, или оно присуще любому хоть сколько-нибудь активному гену? Эксперименты с использованием проб, отображающих весь клеточный пул мРНК, показали, что, хотя, при расщеплении хроматина ДНКазой, среди активных генов существует некоторый разброс по степени уязвимости, приоритет имеют они *все* (даже при очень низком уровне экспрессии, который отображается низкой концентрацией соответствующей мРНК). Гены с неинтенсивной экспрессией, как правило, имеют на себе очень мало рабочих молекул РНК-полимеразы в каждый момент времени; значит, повышенная чувствительность к ДНКазе проистекает не от акта транскрипции, но является отличительной чертой генов, способных (опять же, на данный момент) быть транскрибированными.

Какую протяженность может иметь участок повышенной чувствительности? Это можно определить, используя серию гибридизационных проб, соответствующих активной единице транскрипции с представленными фланкирующими. Участок повышенной чувствительности всегда полностью перекрывает единицу транскрипции (вернее сказать, включает ее в себя). Он может быть фланкирован несколькими тпн ДНК с промежуточной чувствительностью к ДНКазе.

В рассуждении о домене как регуляторной единице есть тот сомнительный аспект, что участок повышенной чувствительности к ДНКазе вообще-то весьма протяженный. В рамках современной концепции любое событие регуляции исследователи склонны связывать с небольшой дискретной областью ДНК (например, инициация транскрипции происходит в промоторе). Даже если это и так, все равно, локализованный тип регуляции должен определять (или подкреплять, или переходить в, или сопровождать, и т.п.) более глобальное и отдаленное изменение организации хроматина как формации, содержащей генетический материал. Этим, кстати, эукариоты отличаются от прокариот.

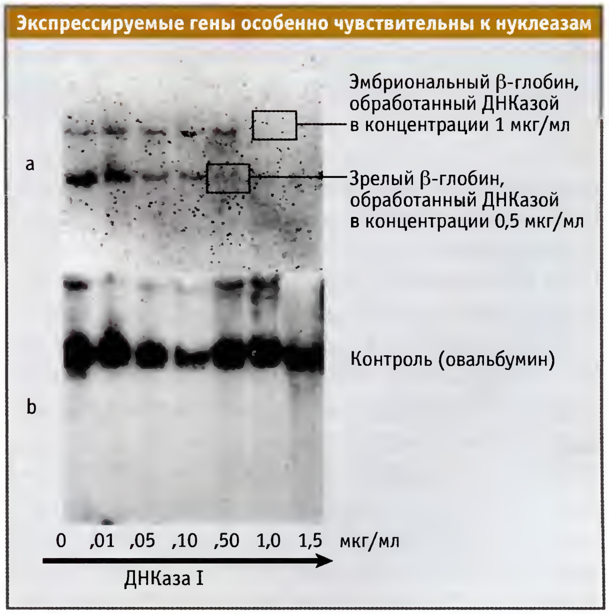


РИС. 29.53 Ген β-глобина, экспрессируемый в эритроидных клетках взрослого организма, очень чувствителен к ДНКазе I. Ген эмбрионального β-глобина тоже чувствителен (но в меньшей степени, и, возможно, из-за своего геномного окружения, поскольку в этих клетках он не экспрессируется). Молчащий ген овальбумина не чувствителен к ДНКазе (в испытанном диапазоне концентраций этого фермента). Фотография любезно предоставлена Harold Weintraub, Fred Hutchinson Cancer Research Center. Использовано с разрешения Mark T. Groudine

29.20 LCR может контролировать домен

Ключевые понятия:

- LCR расположен внутри домена, недалеко от его 5′-конца, и содержит несколько гиперчувствительных сайтов.

Каждый ген контролируется своим промотором, у некоторых есть также энхансеры (содержащие похожие контролирующие элементы, но расположенные далеко); это обсуждалось в гл. 24 «Промоторы и энхансеры». Однако этих элементов контроля бывает недостаточно. Некоторые гены входят в состав домена (вместе с еще несколькими генами), и все они находятся под влиянием регуляторных элементов, воздействующих на домен как целое. Догадка о существовании таких элементов возникла в связи с неспособностью участка ДНК, содержащего ген со всеми его известными регуляторными элементами, должным образом экспрессироваться в качестве трансгена (то есть при искусственном внедрении в геном животного).

Самый яркий пример регулируемого генного кластера — это β-глобины млекопитающих. Здесь нам нужно вспомнить рис. 6.3: α- и β-глобиновые гены млекопитающих существуют как кластеры родственных генов, экспрессирующих на разных сроках развития организма, до и после рождения. Эти гены снабжены большим числом регуляторных элементов, которые подробно изучены. Так, у гена β-глобина, который активен в организме взрослого человека, регуляторные последовательности расположены на обоих флангах (со стороны 5′-конца и со стороны 3′-конца). К этим регуляторным последовательностям относятся элементы позитивного и негативного контроля в области промотора, а также дополнительные элементы позитивного контроля внутри гена и ниже него по ДНК.

Ген β-глобина человека, взятый вместе со всеми регуляторными последовательностями, никогда не станет экспрессироваться в трансгенной мыши даже близко к своей обычной интенсивности. Ему потребуется еще одна регуляторная последовательность. Участки с дополнительной регуляторной функцией можно распознать по сайтам, гиперчувствительным к ДНКазе I и расположенного по краям кластера. Как показано на карте РИС. 29.54, группа гиперчувствительных сайтов расположена на 20 тпн выше гена ε; есть еще одиночный сайт, расположенный на 30 тпн ниже гена β. Выяснилось, что удаление из 5′-концевой области кластера хотя бы одного гиперчувствительного сайта (любого из пяти) значительно снижает уровень экспрессии трансгена в клетках эритролейкемии мыши. Последовательности, лежащие между гиперчувствительными сайтами, могут быть делетированы без особого ущерба для экспрессии трансгена.

5′-концевые гиперчувствительные сайты — это первичные регуляторы, и вся группа их получила название **участок контроля над локусом** (locus control region, LCR). Имеет ли функцию 3′-концевой гиперчувствительный сайт, неясно. Для экспрессии каждого глобинового гена в этом кластере абсолютно необходим LCR. Каждый ген подлежит дальнейшей регуляции своей активности собственными элементами контроля; некоторые из них автономны. Так, экспрессия генов ε и γ, по-видимому, самоуправляема — при условии, что есть LCR. Другие формы контроля основываются на положении гена в кластере, рождая представление о важности *порядка генов* для регуляции их экспрессии.

Весь участок, содержащий глобиновые гены и выступающий далеко за их пределы, представляет собой хромосомный **домен**. Он обладает повышенной чувствительностью к ДНКазе I (см. рис. 29.52). Делеция 5′-концевого LCR снижает чувствительность к ДНКазе всего домена до ее нормального значения, характерного для основной массы неактивного хроматина.

Было предложено две версии работы β-глобинового LCR человека. По одной из версий он помогает активировать промотор, по другой — интенсифицирует транскрипцию с этого промотора. Детали влияния этого LCR на промоторы пока не выяснены.

Применима ли эта модель регуляции к другим генным кластерам? α-глобиновый локус имеет похожую организацию, с гиперчувствительными сайтами, сгруппированными по одну сторону от кластера генов с разными сроками экспрессии; вся область обладает повышенной чувствительностью к ДНКазе. Что касается других случаев контроля LCR над группой генов, то их известно не так много.

В частности, существует LCR, который контролирует гены на более чем одной хромосоме. Этот LCR, T_H2, работает с группой генов, распределенных на участке хромосомы 11 общей протяженностью более 120 тпн. LCR T_H2 координирует активность этих генов, взаимодействуя с их промоторами. Кроме них, LCR T_H2 может взаимодействовать с промотором гена IFNγ на хромосоме 10. Эти взаимодействия взаимно-исключающие, и однозначно взаимосвязаны с судьбой клеток, в которых они происходят.

5′-концевая область глобинового домена содержит LCR

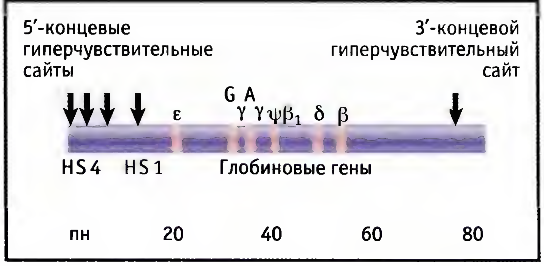


РИС. 29.54 Края глобинового домена отмечены гиперчувствительными сайтами. Группе 5′-концевых гиперчувствительных сайтов соответствует LCR; он имеет принципиальное значение для функции всех генов этого кластера

29.21 Как устроен регуляторный домен?

Ключевые понятия:

- Домен может содержать инсулятор, LCR, сайт прикрепления к ядерному матриксу, и одну или несколько единиц транскрипции.

Обобщая сведения о различных типах регуляторных структур, обнаруживаемых в разных системах, можно, наконец, поразмыслить о том, как устроен хромосомный домен. Отличительной особенностью такого домена является его функциональная ограниченность: регуляторные элементы домена могут действовать только на те единицы транскрипции, которые в нем содержатся. Домен может содержать более чем одну единицу транскрипции и/или энхансер.

На РИС. 29.55 схематично изображены структуры, которые могут участвовать в построении домена.

Инсулятор останавливает активирующие или репрессирующие эффекты на их пути вдоль ДНК. В простейшем случае инсулятор предотвращает распространение регуляторных эффектов, но возможна и более сложная ситуация, когда инсулятор блокирует лишь один тип эффектов и/или действует в одном направлении. Гипотетически, работа инсуляторов затрагивает в хроматине структуру высшего порядка, однако детали и, тем более, разновидности таких событий нам неизвестны.

Участок прикрепления к ядерному матриксу (matrix attachment region, MAR) может отвечать за физическое разделение ДНК на петли, закрепленные на ядерном матриксе (см. разд. 28.6 «В интерфазе ДНК крепится к ядерному матриксу определенными нуклеотидными последовательностями»). Это очень напоминает модель работы инсулятора *gypsy* у *Drosophila*. Действительно, в тестах *in vitro* некоторые последовательности MAR ведут себя как инсуляторы. Но, похоже, их способность прикреплять ДНК к матриксу можно отделить от их способности служить инсуляторами (то есть, эти две способности находятся в компетенции разных элементов последовательности MAR), и простой причинно-следственной механистической связи здесь, скорее всего, нет.



РИС. 29.55 Регуляторные домены являются интегральными частями генома и, в качестве цельных единиц, подлежат разным формам контроля. Инсуляторы предотвращают распространение эффектов между доменами, области MAR обеспечивают прикрепление доменов к ядерному матриксу, а области LCR необходимы для инициации транскрипции во всем домене. Внутри домена есть энхансеры, они могут действовать сразу на несколько промоторов

Скорее всего, эволюционно последовательности некоторых инсуляторов и собственно MAR слились на одном участке ДНК для поддержания единства регуляторной и физической структур хроматина.

Участок LCR действует на расстоянии и может требоваться для экспрессии всех генов домена и каждого из них в отдельности (см. разд. 29.20 «LCR может контролировать домен»). Если в домене есть LCR, от него зависит активность всего домена, но, по всей видимости, наличие в домене LCR — явление, само по себе, достаточно редкое. Существует еще несколько типов последовательностей, способных оказывать *цис*-действие на расстоянии. Как и постулировалось с самого начала, регуляторное свойство LCR сродни свойству энхансера и сосредоточено в гиперчувствительных сайтах, необходимых для полноценной активности промоторов внутри домена.

Организацией регуляторных доменов отчасти объясняется большой размер генома. Домену, как оперативной единице генома, необходимо пространство: например, чтобы хроматин мог в нем деконденсироваться и открыть доступ к ДНК факторам и ферментам. Хотя точная нуклеотидная последовательность большей части домена, очевидно, не важна, увеличение общего количества ДНК внутри домена может поддерживаться естественным отбором; во всяком случае, отбор должен предотвращать слишком близкое расположение различных единиц транскрипции.

29.22 Резюме

Весь эукариотический хроматин состоит из нуклеосом. Одна нуклеосома содержит отрезок ДНК характерной длины (обычно ~200 пн), обернутый вокруг сердцевины — белкового октамера, содержащего по две молекулы гистонов H2A, H2B, H3 и H4. Кроме них, с каждой нуклеосомой ассоциирована одна молекула гистона H1. Практически вся геномная ДНК имеет нуклеосомную организацию; на ее основе в хроматине формируются структуры высших порядков. Обработка массива нуклеосом микрококковой нуклеазой показывает, что его ДНК может быть в рабочем порядке разделена на области двух типов. Связующие области расщепляются нуклеазой быстро; центральные области нуклеосом противятся расщеплению. Диаметр частиц (нуклеосом) определяется тетрамером $H3_2$–$H4_2$. Гистоны H3 и H4 по своей аминокислотной последовательности очень консервативны. Гистоны H2A и H2B менее консервативны; в состав сердцевины нуклеосомы входят два их гетеродимера. Сборка октамеров — это последовательное добавление димеров H2A–H2B к опорному тетрамеру $H3_2$–$H4_2$.

Обходя гистоновый октамер (сердцевину нуклеосомы), ДНК образует –1,65 сверхвитка. ДНК «входит» в нуклеосому и «покидает» ее в близких точках пространства, и может быть «опечатана» по ним гистоном H1. Удале-

ние гистонов сердцевины высвобождает из каждой нуклеосомы по −1,0 сверхвитка ДНК. Разница между −1,0 и −1,65 сверхвитка объясняется тем, что шаг двойной спирали — молекулярная периодичность ДНК — изменяется, когда она переходит из состава нуклеосомы в раствор: от среднего значения в 10,2 пн на виток до 10,5 пн на виток. В пределах нуклеосомы молекулярная периодичность ДНК варьирует от 10,0 пн по краям до 10,7 пн в середине. Траектория ДНК в нуклеосоме имеет изломы.

Массивам нуклеосом свойственно быть уложенными в виде волокон (фибрилл) диаметром 30 нм, устроенных как спиральная обмотка центральной полости (соленоид), содержащая по 6 нуклеосом на виток и упаковывающая ДНК с коэффициетом, равным 40. Удаление гистона H1 приводит к роспуску 30-нм волокна в 10-нм волокно (фибриллу), состоящее из линейной череды нуклеосом. Вероятнее всего, 30-нм волокно устроено как двурядный соленоид, свитый из 10-нм волокна. 30-нм волокно — это главный компонент эухроматина и гетерохроматина; еще более высокие уровни организации хроматина и ультраструктуры митотических хромосом достигаются при участии негистоновых белков.

Сборка нуклеосом может идти двумя путями. В процессе сборки, сопряженном с репликацией, субъединица процессивности реплисомы PCNA рекрутирует белок CAF-1, являющийся фактором сборки нуклеосом. CAF-1 содействует помещению тетрамеров $H3_2$–$H4_2$ на дочерние дуплексы ДНК, появляющиеся в результате репликации. Тетрамеры, которые при этом используются, либо остались от ранее существовавших нуклеосом, разрушенных репликативной вилкой, либо собраны из новосинтезированных гистонов. Димеры H2A–H2B, из тех же источников, монтируются на тетрамер, завершая сборку октамерной сердцевины нуклеосомы. Подбор димеров к тетрамеру происходит случайным образом, так что каждая новая нуклеосома может сочетать в себе «старые» гистоны с новосинтезированными.

Во время транскрипции РНК-полимераза смещает гистоновые октамеры. После того, как полимераза прошла, нуклеосомы снова формируются, если только транскрипция не слишком интенсивна — тогда нуклеосомы могут быть устранены совсем (как в генах рибосомальной РНК). Восстановление на ДНК гистоновых октамеров, смещенных транскрипцией, осуществляет аппарат сборки нуклеосом, не зависящий от репликации. При этом вместо гистона H3 используется гистон H3.3. Похожий способ сборки, не зависящий (по механизму) от репликации и с использованием другой альтернативы H3, применяется для сборки нуклеосом на центромерной ДНК во время ее репликации.

Инсулятор — особая структура в составе хроматина, формирующаяся на определенной последовательности (элементе) геномной ДНК и блокирующая передачу инактивирующих или активирующих влияний между участками ДНК, которые она разделяет. Инсулятор, расположенный между энхансером и промотором, удерживает энхансер от активации промотора. Участок генома, фланкированный инсуляторами, представляет собой обособленный регуляторный домен; эффекты регуляторных взаимодействий внутри такого домена не распространяются за его пределы, и никакие влияния со стороны геномного окружения в домен также не проникают. Большинство инсуляторов блокирует распространение регуляторных эффектов вдоль ДНК в обе стороны, хотя некоторым инсуляторам свойственна проницаемость по одному из направлений. Обычно инсулятор блокирует как активирующие эффекты (взаимодействие энхансера с промотором), так и инактивирующие (опосредованные экспансией гетерохроматина). Полагают, что действие инсуляторов опосредовано изменением высших порядков структуры хроматина, но подробная информация об этом пока отсутствует.

С активностью генов связано изменение чувствительности их ДНК к нуклеазам. Хроматин, в котором возможна транскрипция, отличается повышенной чувствительностью к ДНКазе I. По этому свойству хроматина, проявляющемуся на протяженном участке генома, можно распознать регуляторный домен, содержащий активные или потенциально активные гены. Другой тип изменения чувствительности ДНК к нуклеазам выражается в существовании гиперчувствительных сайтов ДНК, распределенных дискретно. Характерная протяженность гиперчувствительного сайта составляет ~200 пн, присутствие регуляторных белков исключает его нуклеосомную организацию. Гиперчувствительный сайт образует на ДНК границу, которая может более или менее фиксировать позиции нуклеосом на последовательности смежного участка. В общем случае, позиционирование нуклеосом может быть важным фактором контроля доступа к ДНК регуляторных белков.

Гиперчувствительные сайты встречаются в регуляторных последовательностях нескольких типов. Некоторые из этих последовательностей — промоторы, энхансеры, LCR — регулируют транскрипцию. Также гиперчувствительные сайты встречаются в ориджинах репликации, центромерах и ряде других регионов. Промотор или энхансер (в общем случае) действует на один конкретный ген; LCR же содержит целую группу гиперчувствительных сайтов и может держать под контролем регуляторный домен, в котором несколько генов.

Литература

29.2 Любой хроматин состоит из субъединиц – нуклеосом

Обзор

Kornberg, R. D. (1977). Structure of chromatin. *Annu. Rev. Biochem.* 46, 931–954.

McGhee, J. D., and Felsenfeld, G. (1980). Nucleosome structure. *Annu. Rev. Biochem.* 49, 1115–1156.

Исследование

Kornberg, R. D. (1974). Chromatin structure: a repeating unit of histones and DNA. *Science* 184, 868–871.

Richmond, T. J., Finch, J. T., Rushton, B., Rhodes, D., and Klug, A. (1984). Structure of the nucleosome core particle at 7 Å resolution. *Nature* 311, 532–537.

29.3 Геномная ДНК свернута в виде массивов нуклеосом

Исследование

Finch, J. T. et al. (1977). Structure of nucleosome core particles of chromatin. *Nature* 269, 29–36.

29.4 Все нуклеосомы устроены одинаково

Исследование

Shen, X. et al. (1995). Linker histones are not essential and affect chromatin condensation *in vitro*. *Cell* 82, 47–56.

29.5 На поверхности нуклеосомы структура ДНК неоднородна

Обзор

Wang, J. (1982). The path of DNA in the nucleosome. *Cell* 29, 724–726.

Исследование

Richmond, T. J. and Davey, C. A. (2003). The structure of DNA in the nucleosome core. *Nature* 423, 145–150.

29.6 В нуклеосоме молекулярная периодичность ДНК изменяется

Обзор

Travers, A. A. and Klug, A. (1987). The bending of DNA in nucleosomes and its wider implications. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 317, 537–561.

29.7 Организация гистонового октамера

Исследование

Angelov, D., Vitolo, J. M., Mutskov, V., Dimitrov, S., and Hayes, J. J. (2001). Preferential interaction of the core histone tail domains with linker DNA. *Proc. Natl. Acad. Sci. USA* 98, 6599–6604.

Arents, G., Burlingame, R. W., Wang, B.-C., Love, W. E., and Moudrianakis, E. N. (1991). The nucleosomal core histone octamer at 31 Å resolution: a tripartite protein assembly and a left-handed superhelix. *Proc. Natl. Acad. Sci. USA* 88, 10148–10152.

Luger, K. et al. (1997). Crystal structure of the nucleosome core particle at 28 Å resolution. *Nature* 389, 251–260.

29.8 Укладка нуклеосом в волокнах хроматина

Обзор

Felsenfeld, G. and McGhee, J. D. (1986). Structure of the 30 nm chromatin fiber. *Cell* 44, 375–377.

Исследование

Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R. R., and Richmond, T. J. (2004). Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* 306, 1571–1573.

Schalch, T., Duda, S., Sargent, D. F., and Richmond, T. J. (2005). X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* 436, 138–141.

29.9 Воспроизводство хроматина требует сборки нуклеосом

Обзор

Osley, M. A. (1991). The regulation of histone synthesis in the cell cycle. *Annu. Rev. Biochem.* 60, 827–861.

Verreault, A. (2000). De novo nucleosome assembly: new pieces in an old puzzle. *Genes Dev.* 14, 1430–1438.

Исследование

Ahmad, K. and Henikoff, S. (2001). Centromeres are specialized replication domains in heterochromatin. *J. Cell Biol.* 153, 101–110.

Ahmad, K. and Henikoff, S. (2002). The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol. Cell* 9. 1191–1200.

Gruss, C., Wu, J., Koller, T., and Sogo, J. M. (1993). Disruption of the nucleosomes at the replication fork. *EMBO J.* 12, 4533–4545.

Loppin, B., Bonnefoy, E., Anselme, C., Laurencon, A., Karr, T. L., and Couble, P. (2005). The histone H3.3 chaperone HIRA is essential for chromatin assembly in the male pronucleus. *Nature* 437, 1386–1390.

Ray-Gallet, D., Quivy, J. P., Scamps, C., Martini, E. M., Lipinski, M., and Almouzni, G. (2002). HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. *Mol. Cell* 9, 1091–1100.

Shibahara, K., and Stillman, B. (1999). Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. *Cell* 96, 575–585.

Smith, S. and Stillman, B. (1989). Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication *in vitro*. *Cell* 58, 15–25.

Smith, S. and Stillman, B. (1991). Stepwise assembly of chromatin during DNA replication *in vitro*. *EMBO J.* 10, 971–980.

Tagami, H., Ray-Gallet, D., Almouzni, G., and Nakatani, Y. (2004). Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* 116, 51–61.

Yu, L. and Gorovsky, M. A. (1997). Constitutive expression, not a particular primary sequence, is the important feature of the H3 replacement variant hv2 in *Tetrahymena thermophila*. *Mol. Cell. Biol.* 17, 6303–6310.

29.11 Организованы ли транскрибируемые гены в нуклеосомы?

Обзор

Kornberg, R. D. and Lorch, Y. (1992). Chromatin structure and transcription. *Annu. Rev. Cell Biol.* 8, 563–587.

29.12 Гистоновые октамеры смещаются транскрипцией

Исследование

Cavalli, G. and Thoma, F. (1993). Chromatin transitions during activation and repression of galactose-regulated genes in yeast. *EMBO J.* 12, 4603–4613.

Studitsky, V. M., Clark, D. J., and Felsenfeld, G. (1994). A histone octamer can step around a transcribing polymerase without leaving the template. *Cell* 76, 371–382.

29.13 Смещение нуклеосом и формирование их заново требует участия специальных факторов

Исследование

Belotserkovskaya, R., Oh, S., Bondarenko, V. A., Orphanides, G., Studitsky, V. M., and Reinberg, D. (2003). FACT facilitates transcription-dependent nucleosome alteration. *Science* 301, 1090–1093.

Saunders, A., Werner, J., Andrulis, E. D., Nakayama, T., Hirose, S., Reinberg, D., and Lis, J. T. (2003). Tracking FACT and the PHK polymerase II elongation complex through chromatin *in vivo*. *Science* 301, 1094–1096.

29.14 Инсуляторы блокируют действие энхансеров и гетерохроматина

Обзор

Gerasimova, T. I. and Corces, V. G. (2001). Chromatin insulators and boundaries: effects on transcription and nuclear organization. *Annu. Rev. Genet.* 35, 193–208.

West, A. G., Gaszner, M., and Felsenfeld, G. (2002). Insulators: many functions, many mechanisms. *Genes Dev.* 16, 271–288.

29.15 Инсуляторы могут ограничивать домен

Исследование

Chung, J. H., Whiteley, M., and Felsenfeld, G. (1993). A 5′ element of the chicken b-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell* 74, 505–514.

Cuvier, O., Hart, C. M., and Laemmli, U. K. (1998). Identification of a class of chromatin boundary elements. *Mol. Cell Biol.* 18, 7478–7486.

Gaszner, M., Vazquez, J., and Schedl, P. (1999). The Zw5 protein, a component of the scs chromatin domain boundary, is able to block enhancer-promoter interaction. *Genes Dev.* 13, 2098–2107.

Kellum, R. and Schedl, P. (1991). A position-effect assay for boundaries of higher order chromosomal domains. *Cell* 64, 941–950.

Pikaart, M. J., Recillas-Targa, F., and Felsenfeld, G. (1998). Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. *Genes Dev.* 12, 2852–2862.

Zhao, K., Hart, C. M., and Laemmli, U. K. (1995). Visualization of chromosomal domains with boundary element-associated factor BEAF-32. *Cell* 81, 879–889.

29.16 Инсулятор транспозона *gypsy* и его белковый аппарат

Исследование

Gerasimova, T. I., Byrd, K., and Corces, V. G. (2000). A chromatin insulator determines the nuclear localization of DNA. *Mol. Cell* 6, 1025–1035.

Harrison, D. A., Gdula, D. A., Cyne, R. S., and Corces, V. G. (1993). A leucine zipper domain of the suppressor of hairy-wing protein mediates its repressive effect on enhancer function. *Genes Dev.* 7, 1966–1978.

Roseman, R. R., Pirrotta, V., and Geyer, P. K. (1993). The su(Hw) protein insulates expression of the *D. melanogaster white* gene from chromosomal position-effects. *EMBO J.* 12, 435–442.

29.17 Инсуляторы бывают разными по силе

Исследование

Hagstrom, K., Muller, M., and Schedl, P. (1996). Fab-7 functions as a chromatin domain boundary to ensure proper segment specification by the *Drosophila* bithorax complex. *Genes Dev.* 10, 3202–3215.

Mihaly, J. et al. (1997). *In situ* dissection of the Fab-7 region of the bithorax complex into a chromatin domain boundary and a polycomb-response element. *Development* 124, 1809–1820.

Zhou, J. and Levine, M. (1999). A novel *cis*-regulatory element, the PTS, mediates an anti-insulator activity in the *Drosophila* embryo. *Cell* 99, 567–575.

29.18 Сайты, гиперчувствительные к ДНКазе, реагируют на изменения в структуре хроматина

Обзор

Gross, D. S. and Garrard, W. T. (1988). Nuclease hypersensitive sites in chromatin. *Annu. Rev. Biochem.* 57, 159–197.

Исследование

Groudine, M., and Weintraub, H. (1982). Propagation of globin DNAase I-hypersensitive sites in absence of factors required for induction: a possible mechanism for determination. *Cell* 30, 131–139.

Moyne, G., Harper, F., Saragosti, S., and Yaniv, M. (1982). Absence of nucleosomes in a histone-containing nucleoprotein complex obtained by dissociation of purified SV40 virions. *Cell* 30, 123–130.

Scott, W. A. and Wigmore, D. J. (1978). Sites in SV40 chromatin which are preferentially cleaved by endonucleases. *Cell* 15, 1511–1518.

Varshavsky, A. J., Sundin, O., and Bohn, M. J. (1978). SV40 viral minichromosome: preferential exposure of the origin of replication as probed by restriction endonucleases. *Nucleic Acids Res.* 5, 3469–3479.

29.19 Участки хроматина, содержащие активные гены, обладают повышенной чувствительностью к нуклеазам

Исследование

Stalder, J. et al. (1980). Tissue-specific DNA cleavage in the globin chromatin domain introduced by DNAase I. *Cell* 20, 451–460.

29.20 LCR может контролировать домен

Обзор

Bulger, M. and Groudine, M. (1999). Looping versus linking: toward a model for long-distance gene activation. *Genes Dev.* 13, 2465–2477.

Grosveld, F., Antoniou, M., Berry, M., De Boer, E., Dillon, N., Ellis, J., Fraser, P., Hanscombe, O., Hurst, J., and Imam, A. (1993). The regulation of human globin gene switching. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 339, 183–191.

Исследование

Gribnau, J., de Boer, E., Trimborn, T., Wijgerde, M., Milot, E., Grosveld, F., and Fraser, P. (1998). Chromatin interaction mechanism of transcriptional control *in vitro*. *EMBO J.* 17, 6020–6027.

Spilianakis, C. G., Lalioti, M. D., Town, T., Lee, G. R., and Flavell, R. A. (2005). Interchromosomal associations between altePHKtively expressed loci. *Nature* 435, 637–645.

van Assendelft, G. B., Hanscombe, O., Grosveld, F., and Greaves, D. R. (1989). The b-globin dominant control region activates homologous and heterologous promoters in a tissue-specific manner. *Cell* 56, 969–977.

29.21 Как устроен регуляторный домен?

Обзор

West, A. G., Gaszner, M., and Felsenfeld, G. (2002). Insulators: many functions, many mechanisms. *Genes Dev.* 16, 271–288.

Контроль структуры хроматина

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

30.1 Введение

Если рассматривать транскрипцию с точки зрения молекулярных взаимодействий, в которых участвуют ДНК, факторы транскрипции и РНК-полимераза, то, в идеале, можно получить аккуратное описание того, что происходит *in vitro*, но упустить одну важнейшую особенность, присущую событиям *in vivo*. Особенность состоит в том, что реальный геном имеет нуклеосомную организацию, а транскрипция никак не может быть инициирована на промоторном участке ДНК, упакованном в нуклеосомы. В некотором смысле, гистоны функционируют как общие репрессоры транскрипции: идея не новая, хотя в этой главе мы увидим, что гистоны участвуют и в более специфических взаимодействиях. Активация гена требует изменения состояния хроматина: очень важно понять, каким образом факторам транскрипции удается получить доступ к промоторной ДНК.

Локальная структура хроматина — это неотъемлемая часть системы контроля экспрессии генов. Каждый ген может пребывать в одном из двух состояний (статусов экспрессии). В активном состоянии ген можно застать только в клетках, где он экспрессируется. Изменение структуры гена предшествует акту его транскрипции и показывает, что ген, в принципе, транскрибируем. Это означает, что активация гена должна быть первым этапом его экспрессии. Активные гены можно обнаружить в доменах эухроматина с повышенной чувствительностью к нуклеазам (см. разд. 29.19 «Участки хроматина, содержащие активные гены, обладают повышенной чувствительностью к нуклеазам»). Сайты, гиперчувствительные к нуклеазам, возникают в промоторах перед тем, как ген активируется (см. разд. 29.18 «Сайты, гиперчувствительные к ДНКазе, реагируют на изменения в структуре хроматина»).

Относительно недавно выяснилось, что между инициацией транскрипции и структурой хроматина существует глубокая связь. Некоторые активаторы транскрипции напрямую модифицируют гистоны; в частности, с активацией генов связано ацетилирование гистонов. Некоторые репрессоры транскрипции, наоборот, деацетилируют гистоны. Так обратимые модификации гистонов в области промотора служат инструментом контроля экспрессии гена и частью механизма, поддерживающего ген в активном или неактивном состоянии.

Механизмы локального поддержания хроматина в неактивном состоянии взаимосвязаны со средствами репрессии конкретного промотора. Участие негистоновых белков в создании гетерохроматина опосредовано все теми же гистонами; модификация гистонов может быть важна для этих взаимодействий. Изменения в структуре хроматина, однажды установившись и перманентно стабилизировавшись относительно смены фаз клеточного цикла (в том числе, делений), могут обусловить **эпигенетическое** влияние, то есть наследуемое изменение в активности гена без изменения последовательности его ДНК. При этом особенности функционирования гена определяются самоподдерживающейся структурой хроматина. В самом выборе термина «эпигенетический» (*над*-генетический) отражен тот факт, что у гена может быть унаследованное состояние (активное или неактивное), не зависящее от его последовательности. Примером крайней степени проявления эпигенеза служат самоорганизующиеся структуры **прионов** (инфекционных агентов белковой природы).

30.2 Хроматин может находиться в альтернативных состояниях

Ключевые понятия:

- Структура хроматина стабильна и не может быть преобразована изменением равновесия между факторами транскрипции и гистонами.

Для того чтобы объяснить, как меняется статус экспрессии гена, было предложено два типа моделей: равновесия и резкой смены состояния.

На рис. 30.1 показана модель равновесия. В ней важна только концентрация активатора либо репрессора, которой задается равновесие между его свободной формой и формой, связанной с ДНК. Активация либо репрессия идет путем связывания с конкретной последовательностью-мишенью (или несколькими такими последовательностями). Когда концентрация белкового фактора (активатора либо репрессора) достаточно высока, он обязательно оккупирует последовательность-мишень и влияет на статус экспрессии гена. Этот тип моделей хорошо объясняет регуляцию транскрипции в бактериальных клетках, где экспрессия генов всецело определяется действиями конкретных белков — репрессоров и активаторов (см. гл. 12 «Оперон»). Транскрибируется бактериальный ген или нет, можно предсказать, оперируя числовыми значениями концентраций различных факторов, активирующих либо репрессирующих конкретный ген. При изменении этих концентраций — *в любой момент времени* — статус экспрессии изменится соответственно. В большинстве случаев связывание молекул фактора с ДНК идет кооперативно, так что, как только концентрация фактора стала достаточно высокой, происходит его быстрая ассоциация с ДНК, приводя к переключению статуса экспрессии гена.

С эукариотическим хроматином дело обстоит иначе. В классических экспериментах *in vitro* было показано, что активное либо неактивное состояние может, установившись, не поддаваться последующему воздействию добавляемых компонентов. Фактор транскрипции $TF_{III}A$, необходимый для того, чтобы РНК-полимераза III транскрибировала гены 5S рРНК, не может активировать свои гены мишени *in vitro*, если они состоят в комплексе с гистонами. В присутствие свободной ДНК фактор сформиру-



Рис. 30.1 Согласно принципу равновесия, занятость сайта на ДНК зависит от концентрации белка, способного связаться с этим сайтом

ет на ней комплекс транскрипции, и последующее добавление гистонов уже не помешает: ген останется активным. Если фактор связался со своим сайтом, то никуда оттуда не денется; это позволит череде молекул РНК-полимеразы инициировать транскрипцию. Для успеха транскрипции критично, какой белок — гистоны или фактор — окажется на ДНК *первым*.

На Рис. 30.2 изображены два типа условий, которые могут сложиться в эукариотическом промоторе. Если ген не активен, на промоторе присутствуют нуклеосомы, удерживающие ДНК от связывания с фактором транскрипции и РНК-полимеразой. Когда ген активен, его промотор оккупирован базовым аппаратом транскрипции, а гистоновые октамеры диссоциированы. Любое из этих двух состояний промотора будет стабильным. Похожая ситуация вырисовывается с комплексом $TF_{II}D$ на промоторах для РНК-полимеразы II. Плазмида, содержащая аденовирусный промотор, может быть транскрибирована РНК-полимеразой II *in vitro*; реакция требует участия $TF_{II}D$ и других факторов транскрипции. Матрицу можно упаковать в нуклеосомы, добавив в реакцию гистоны. Если добавить сначала гистоны, а потом $TF_{II}D$, транскрипция инициироваться не сможет. Если сначала добавить $TF_{II}D$, то матрица может быть транскрибирована прямо в составе экстрагированного хроматина. $TF_{II}D$, таким образом, узнает свободную ДНК; нуклеосомную ДНК он либо не узнает, либо не может с ней работать. Для инициации транскрипции достаточно, чтобы $TF_{II}D$ встретился с ДНК до прихода гистонов; другие факторы транскрипции и РНК-полимераза могут быть добавлены потом. По-видимому, связывание $TF_{II}D$ с промотором закладывает структурную основу, с которой уже могут связаться остальные компоненты аппарата транскрипции.

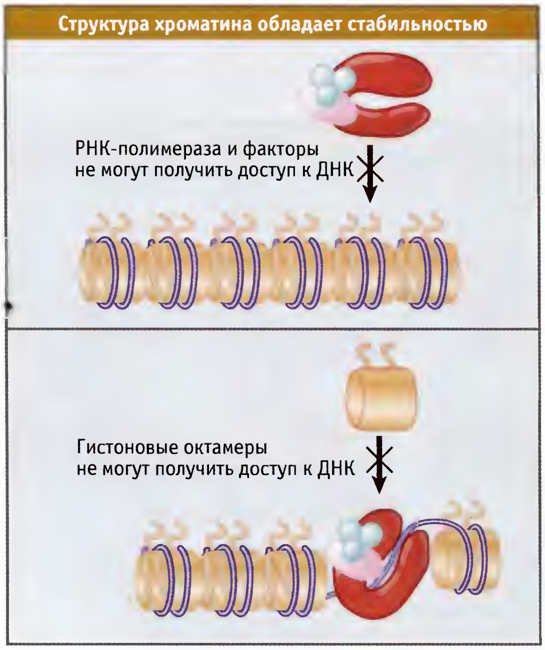


Рис. 30.2 Факторы транскрипции (и РНК-полимераза) не могут связаться с промотором, который скрыт в нуклеосомах. Если же факторы транскрипции (и РНК-полимераза) связались с каким-либо промотором и образовали на нем стабильный комплекс инициации, этот промотор уже нельзя скрыть в нуклеосомах

Отметим, что ситуацию в системах *in vitro* нельзя приравнивать к естественной, поскольку там используются другие пропорции компонентов, чем присутствуют *in vivo*. Главный вывод из полученных результатов в том, что *факторы транскрипции или нуклеосомы могут образовывать стабильные структуры, которые невозможно изменить простым сдвигом баланса свободных компонентов.*

30.3 Ремоделирование хроматина – это активный процесс

Ключевые понятия:

- Существуют белковые комплексы ремоделирования хроматина; их энергоснабжение идет за счет гидролиза АТФ.
- Комплексы SWI/SNF, RSC и NURF довольно крупные; некоторые субъединицы у всех трех комплектов очень похожи.
- У комплексов ремоделирования нет специфичности по отношению к конкретным сайтам на ДНК; к работе на промоторе такой комплекс привлекается компонентом аппарата транскрипции

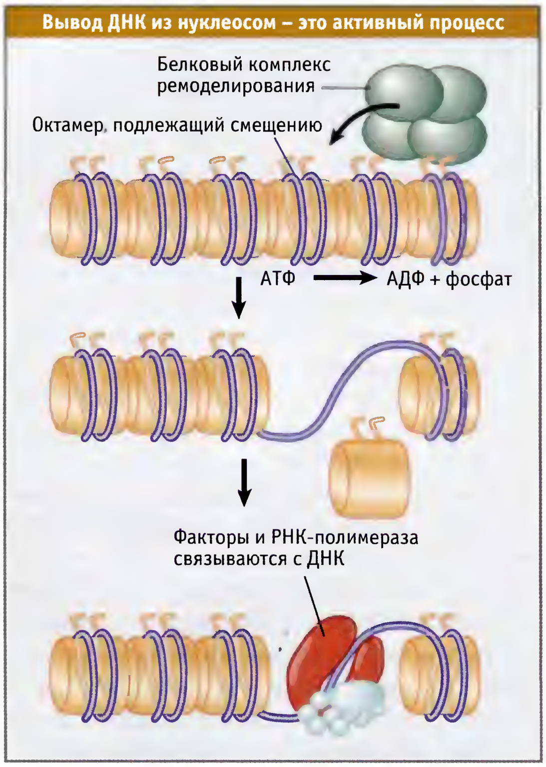


РИС. 30.3 Динамическая модель «транскрипции хроматина» основана на участии белковых факторов, которые, используя энергию гидролиза АТФ, нарушают нуклеосомную организацию конкретных последовательностей ДНК

Процесс наведения изменений в структуре хроматина называют **ремоделированием хроматина**. В это понятие входит смещение гистонов, требующее энергозатрат. Чтобы высвободить гистоны из хроматина, необходимо разрушить множество контактов белков между собой и с ДНК, для прекращения этих контактов нужна энергия. На РИС. 30.3 изображен принцип динамической модели (*dynamic model*) смещения нуклеосом неким фактором, гидролизующим АТФ. Когда гистоновый октамер покидает ДНК, с ней могут связаться другие белки (в данном случае, факторы транскрипции и РНК-полимераза).

На РИС. 30.4 приведены схемы трех типов ремоделирования, которые могут быть воспроизведены и охарактеризованы *in vitro*:

- Гистоновые октамеры могут скользить вдоль ДНК, со сдвигом нуклеиновой кислоты относительно массива белков. Это меняет расположение конкретных последовательностей ДНК на поверхности нуклеосом.
- Просветы между гистоновыми октамерами могут меняться, опять же с тем результатом, что позиции конкретных последовательностей ДНК относительно белков становятся другими.
- Самое экстенсивное ремоделирование — это полное смещение одного или нескольких октамеров с ДНК, с образованием пробела в нуклеосомной организации (то есть участка ДНК, свободного от нуклеосом).

Ремоделирование хроматина чаще всего применяется в промоторах генов, предназначенных для транскрипции. Оно, как правило, необходимо, чтобы аппарат транскрипции смог получить доступ к промотору. Ремоделирование хроматина может потребоваться и в других случаях, например при репарации поврежденной ДНК.

Самый распространенный вариант ремоделирования — это смещение с ДНК одного или нескольких гистоновых октамеров. Что такое событие произошло, можно заметить по изменению паттерна «лестницы» на электрофореграмме ДНК (той области, где она лишилась нуклеосомной «протекции») после расщепления хроматина микрококковой нуклеазой. В пробеле свободной ДНК, который при этом возникает посреди массива нуклеосом, нередко можно обнаружить сайт, гиперчувствительный к ДНКазе I (см. разд. 29.18 «Сайты, гиперчувствительные к ДНКазе, как индикаторы изменений в структуре хроматина»).

Иногда ремоделирование хроматина не столь «драматично» и заключается, например, в сдвиге ротационного позиционирования одной-единственной нуклеосомы. Это изменение можно засечь по «размыванию лестницы» с шагом в 10 оснований, обычно возникающей при обработке нуклеосомы ДНКазой I. Таким образом, спектр изменений структуры хроматина (в рамках ремоделирования) достаточно широк: от небольшого сдвига в расположении нуклеосом(ы) до полного исключения нуклеосом с участка ДНК.

Ремоделированием хроматина занимаются большие белковые комплексы, использующие гидролиз АТФ для обеспечения необходимой энергии. АТФазная субъеди-

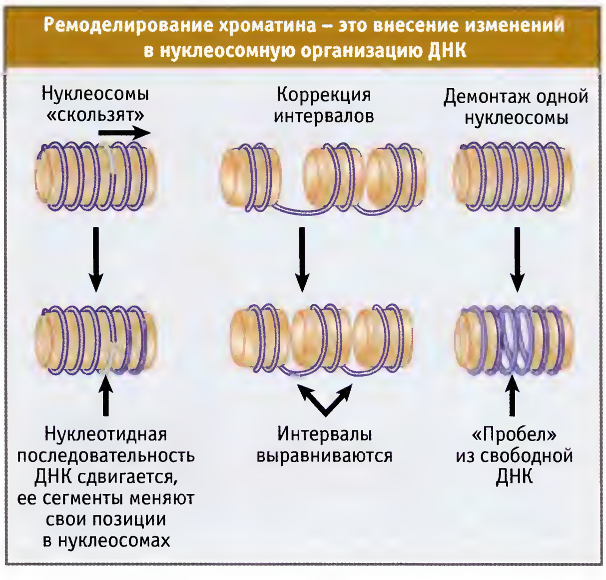


РИС. 30.4 Комплексы ремоделирования могут заставить нуклеосомы «скользить» вдоль ДНК. Они также могут демонтировать отдельные нуклеосомы или корректировать размер межнуклеосомных интервалов (связующих сегментов ДНК)

ница — это, во всех смыслах, центральная часть такого комплекса. Комплексы ремоделирования хроматина классифицируют по типу АТФазной субъединицы: комплексы, у которых эти субъединицы похожи, относят к одному семейству. Обычно у таких «родственных» комплексов и некоторые другие субъединицы тоже общие. Из таблицы на РИС. 30.5 явствует, что два основных типа комплексов — это SWI/SNF и ISWI (что означает «имитация SWI»). В дрожжевых клетках присутствует две разновидности комплекса SWI/SNF и три разновидности комплекса ISWI. Комплексы обоих типов есть также у мух и у людей. Каждому типу комплексов присущ свой набор ремоделирующих активностей.

Комплекс **SWI/SNF** открыли первым. Его назвали «SWI/SNF», потому что многие из его субъединиц кодируются генами, которые у дрожжей *S. cerevisiae* прежде идентифицировались в качестве носителей мутаций SWI или SNF. Эти мутации плейотропны и по набору вызываемых дефектов похожи на мутантов, у которых утрачен С-концевой домен (или «С-концевой хвост») РНК-полимеразы II. Кроме того, мутации в SWI и SNF генетически взаимодействуют (супрессируются)мутациями в генах, кодирующих компоненты хроматина, а именно *SIN1* (кодирует негистоновый белок) и *SIN2* (кодирует гистон Н3). Продукты генов *SWI* и *SNF* требуются для экспрессии генов во множестве локусов (всего, при их инактивации, страдает экспрессия ~120, или 2%, генов *S. cerevisiae*), где комплекс SWI/SNF востребован как участник ремоделирования хроматина на промоторе или в его окрестностях.

Комплекс SWI/SNF каталитически активен *in vitro*; в дрожжевой клетке насчитывается всего лишь ~150 таких комплексов. Каждый из генов, кодирующих субъединицы SWI/SNF, сам по себе не принципиален для выживания клетки; что говорит о существовании других путей ремоделирования хроматина у дрожжей. Комплекс RSC (remodels the structure of chromatin; «ремоделирует структуру хроматина») представлен у тех же дрожжей обильнее, чем SWI/SNF. Комплекс RSC действует в ~700 локусах-мишенях. Отсутствие в клетке любого из описанных комплексов летально.

Комплексы ремоделирования бывают нескольких типов			
Тип комплекса	SWI/SNF	ISWI	Другие
Дрожжи	SWI/SNF RSC	ISW1 ISW2	INO80 complex SWRI
Муха	dSWI/SNF (brahma)	NURF CHRAC ACF	
Человек	hSWI/SNF	RSF hACF/WCFR hCHRAC WICH	NuRD INO80 complex SRCAP
Лягушка		WICH CHRAC ACF	Mi-2

РИС. 30.5 Комплексы ремоделирования можно классифицировать по строению АТФазной субъединицы

Комплексы SWI/SNF способны ремоделировать хроматин *in vitro* либо без потери гистонов, либо смещая с ДНК гистоновые октамеры. В реакциях обоих типов изменение структуры октамера проходит через один и тот же промежуточный этап, который может вылиться либо в переоформление нуклеосомы на той же самой молекуле ДНК, либо в смещение гистонового октамера на другую молекулу ДНК. Комплекс SWI/SNF индуцирует изменения в контактах белок-ДНК, что сопровождается изменениями чувствительности к ДНКазе I. Эти изменения остаются в силе и после того, как комплекс покинет нуклеосому. Субъединица Swi2 — это АТФаза, снабжающая комплекс энергией.

Между ДНК и гистоновым октамером в любой нуклеосоме существует множество контактов. При анализе кристаллической структуры таких контактов обнаруживается четырнадцать. Чтобы октамер мог освободиться или передвинуться в другое положение, все эти контакты необходимо разорвать. Как же этого добиться? Некоторые очевидные механизмы следует исключить из рассмотрения, потому что мы знаем, что во время ремоделирования хроматина одноцепочечная ДНК не образуется, и никакой хеликазной активности с комплексами ремоделирования хроматина не ассоциировано. Современная точка зрения состоит в том, что ремоделирующие комплексы семейств SWI/SNF и ISWI используют гидролиз АТФ для вращения ДНК на поверхности нуклеосомы. Есть косвенные свидетельства того, что это создает механическую силу, помогающую небольшому сегменту ДНК высвободиться с поверхности нуклеосомы и поменять свое положение.

Одна из важных реакций, проводимых комплексами ремоделирования, создает эффект скольжения нуклеосом. Было установлено, что комплексы семейства ISWI могут изменять позиционирование нуклеосом без отрыва гистоновых октамеров от ДНК: они движутся вдоль нее, благодаря чему реакцию назвали реакцией скольжения. Скольжение можно запретить, удалив у гистонов Н4 их N-концевые «хвосты» (но функция «хвостов» в данном случае доподлинно не известна). Комплексы SWI/SNF тоже могут заставить нуклеосому поменять положение относительно последовательности ДНК; реакцию можно предотвратить внесением в ДНК барьера, что свидетельствует об истинном скольжении, при котором гистоновый октамер движется по ДНК более или менее равномерно, не теряя с ней контакта.

Одной из загадок в работе SWI/SNF являются размеры комплекса. В нем одиннадцать субъединиц, общей молекулярной массой ~2 × 10^6. РНК-полимераза и нуклеосомы откровенно мелки по сравнению со SWI/SNF, и трудно себе представить, каким образом множеству компонентов комплекса удается взаимодействовать с ДНК, зафиксированной на поверхности нуклеосомы. Полностью активный комплекс транскрипции, названный «холоферментом РНК-полимеразы II», содержит саму полимеразу, все факторы TF_{II} (кроме ТВР и $TF_{II}A$),

а также комплекс SWI/SNF, ассоциированный с С-концевым хвостом полимеразы. На практике, весь клеточный запас комплекса SWI/SNF можно экстрагировать в составе препаратов холофермента. Это, в частности, означает, что ремоделирование хроматина и узнавание промоторов транскрипции производятся единым белковым комплексом.

30.4 Нуклеосомная организация промотора может быть нестандартной

Ключевые понятия:

- **Привлечением комплексов ремоделирования к работе над изменением структуры хроматина занимаются активаторы транскрипции, которые специфичны по отношению к промоторам.**
- **Как только ремоделирующий комплекс связался с промотором, фактор, привлекший его туда, может быть свободен.**
- **Для того чтобы активатор транскрипции смог связаться с ДНК в промоторе MMTV, требуется изменить ротационное позиционирование одной из нуклеосом.**

Как происходит нацеливание белковых комплексов ремоделирования хроматина на конкретные сайты? Сами комплексы не содержат субъединиц, специфичных к конкретным последовательностям ДНК. В голову приходит модель, изображенная на РИС. 30.6: чтобы комплекс ремоделирования работал на промоторе, его рекрутируют (физически привлекают туда) другие белки — активаторы либо, что тоже бывает, репрессоры транскрипции.

Участие комплексов ремоделирования хроматина в активации генов было открыто через изучение факторов транскрипции, нуждающихся в этих комплексах, чтобы активировать свои гены-мишени. Одним из первых примеров был фактор GAGA, активирующий *in vitro* промотор *hsp70 D. melanogaster*. Связывание GAGA с четырьмя $(CT)_n$-богатыми сайтами промотора видоизменяет нуклеосомную организацию, создает гиперчувствительный сайт и позиционирует нуклеосомы на смежных участках. Вмешательство в нуклеосомную организацию — это энергозависимый процесс, который был бы невозможен без участия ремоделирующего комплекса NURF. В результате деятельности NURF в последовательности ДНК возникает граница, задающая позиции нуклеосом на смежных участках. Тем временем фактор GAGA связывает ДНК по своим сайтам-мишеням; присутствие на ДНК фактора GAGA фиксирует хроматин в ремоделированном состоянии.

Взаимодействие комплексов ремоделирования с факторами транскрипции — это ключ к пониманию их *modus operandi*. У дрожжей фактор Swi5 активирует локус *HO*. Заметим, что Swi5 не входит в комплекс SWI/SNF. В конце митоза фактор Swi5 попадает в ядро и связывает там промотор *HO*. Затем Swi5 рекрутирует SWI/SNF и покидает промотор, оставив на нем SWI/SNF. Это означает, что фактор транскрипции может активировать промотор по принципу «ударь и беги», согласно которому фактор выполнил свою работу, как только комплекс ремоделирования хроматина связался с промотором.

Одним из первых регуляторных контуров, для которых было показано, что в активацию гена вовлечено изменение в нуклеосомной организации его промотора, была система *PHO*. С промотором *PHO5* ассоциирован фактор типа bHLH (домен со свойствами основания, примыкающий к структуре «спираль-петля-спираль»), называемый PHO4. Фактор PHO4, в ответ на нехватку фосфата, индуцирует устранение с промотора *PHO5* четырех точно позиционированных нуклеосом. Этот эффект независим от транскрипции (поскольку точно так же происходит у мутантов *TATA*$^-$) и независим от реп-

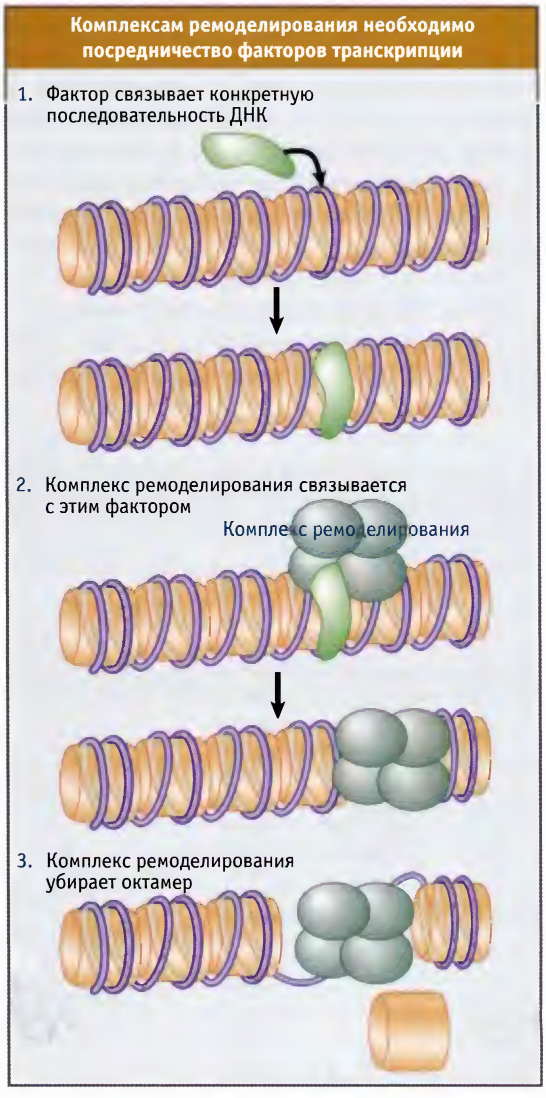


РИС. 30.6 Комплекс ремоделирования связывается с хроматином через активатор (или через репрессор)

ликации. В промоторе *PHO5* есть два сайта связывания PHO4. Один из них расположен между нуклеосомами и может быть связан изолированным ДНК-связывающим доменом PHO4. Другой сайт находится в нуклеосоме, где PHO4 его не узнает. Расщепление этой нуклеосомы, дающее PHO4 возможность связаться со вторым своим сайтом, необходимо для активации гена. Чтобы второе связывание произошло, требуется присутствие другого, активирующего транскрипцию, домена PHO4. Домен-активатор в белке PHO4 можно заменить аналогичным доменом другого фактора транскрипции, VP16: развал нуклеосом пройдет с той же специфичностью. Это означает, что прицельный развал нуклеосомной организации на промоторе *PHO5* обеспечивается белок-белковыми взаимодействиями, в которых участвует тот же домен фактора, который затем активирует транскрипцию: тоже через образование белок-белковых контактов, но уже не с комплексом ремоделирования хроматина, а с собственно аппаратом транскрипции. Какой именно из комплексов ремоделирования занимается этим промотором, не известно.

Масштабное обследование дрожжевого генома на предмет расположения нуклеосом показало, что большинство сайтов связывания факторов транскрипции от нуклеосом свободно. Промоторы для РНК-полимеразы II, как правило, имеют участок ~200 пн, свободный от нуклеосом и находящийся выше точки старта транскрипции. Этот участок фланкирован позиционированными нуклеосомами.

Инициация транскрипции не обязательно должна сопровождаться исключением нуклеосом. Некоторые активаторы могут связать ДНК на поверхности нуклеосомы. Элементы HRE (сокращение от английского "hormone response element") — это участки ДНК, присутствующие в промоторах некоторых генов: через них идет клеточный ответ на стимуляцию (стероидным) гормоном. В некоторых HRE нуклеосомы позиционированы со всей возможной точностью, так что рецепторы стероидных гормонов имеют гарантированный доступ к сайтам связывания. Связывание рецептора с ДНК может повлиять на ее взаимодействие с гистонами и даже повлечь за собой открывание новых сайтов связывания. Точное позиционирование нуклеосом в этом случае может требоваться либо для «презентации» ДНК в конкретной ротационной фазе, либо для белок-белковых взаимодействий между активаторами и гистонами (или другими компонентами хроматина). Таким образом, мы уже немного отошли от видения хроматина исключительно как репрессивной структуры и приблизились к рассмотрению хроматина как интерактивной сферы действия активаторов транскрипции.

Один из промоторов, нуждающихся в специфической нуклеосомной организации, — это промотор MMTV вируса. Нуклеотидная последовательность промотора MMTV содержит массив из шести частично палиндромных сайтов, каждый из которых связывается одним димером рецептора гормона (hormone receptor, HR). Все вместе эти сайты составляют элемент HRE. Кроме того, в промоторе MMTV есть одиночный сайт связывания фактора NF1, а по соседству — два сайта для фактора OTF. Рецептор гормона HR и фактор NF1 не могут одновременно связаться со своими сайтами на свободной (безнуклеосомной) ДНК. На РИС. 30.7 показано, как при помощи нуклеосомной организации контролируется связывание факторов с промотором MMTV.

Будучи активирован гормоном, HR покрывает свои сайты, но не обеспечивает защиты промотора от микрококковой нуклеазы (сайты ДНК, чувствительные к этому ферменту, расположены по краям нуклеосомы). Это говорит о том, что HR связывается с ДНК на поверхности нуклеосомы. Однако до проведения стимуляции гормоном ротационное позиционирование ДНК на поверхности нуклеосомы дозволяет доступ лишь к двум сайтам связывания HR. Чтобы димеры HR могли связать в HRE еще два сайта, требуется изменить ротационное позиционирование этой нуклеосомы. Изменение можно детектировать по появлению сайта, чувствительного к нуклеазам, ровно по центру нуклеосомы (как раз посередине элемента HRE). Фактор NF1 (footprint) на нуклеосоме только после стимуляции гормоном. По всей видимости, структурные изменения, обусловленные связыванием HR, способствуют связыванию NF1. Отчасти этот эффект может объясняться снятием стерического препятствия, если HR создает загромождение и препятствует связыванию NF1 со свободной ДНК, но не с ДНК в составе нуклеосомы.

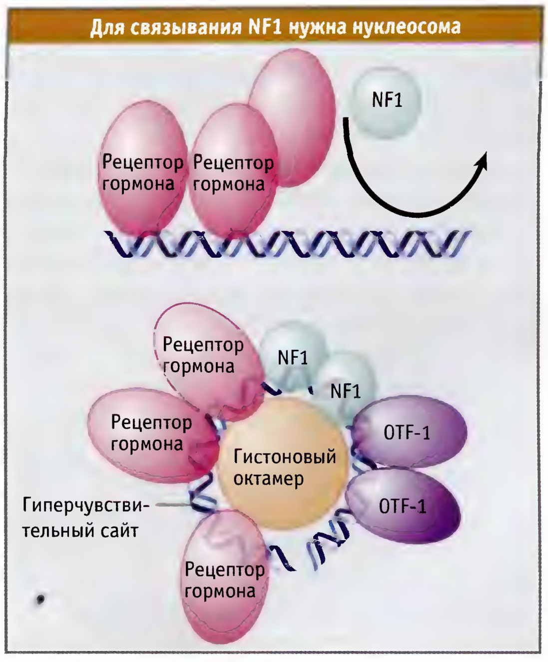


РИС. 30.7 Рецептор гормона и фактор транскрипции NF1 не могут одновременно связать промотор MMTV, если он находится в форме линейной ДНК без вторичной структуры. Нуклеосомная организация ДНК промотора MMTV решает эту проблему

30.5 Модификация гистонов как ключевой фактор изменения структуры хроматина

Ключевые понятия:

- Гистоны подлежат модификации метилированием, ацетилированием или фосфорилированием.

Экспрессируется ген или нет, зависит от структуры хроматина и в области промотора, и в окружающем хромосомном домене. Структура хроматина может регулироваться как отдельными событиями активации, так и изменениями, затрагивающими обширную область хромосомы. Самые локальные события изменения структуры хроматина касаются лишь отдельно взятого гена-мишени, точнее, нуклеосомной организации его промотора и/или ближайших окрестностей промотора. В других случаях изменения могут затронуть более обширную область, вплоть до целой хромосомы.

Структурные изменения, охватывающие обширную область генома, контролируют потенциальную возможность экспрессии генов в этой области. Термином **умолкание** (сайленсинг) пользуются, чтобы обозначить репрессию активности генов в некоторой локальной области хромосомы. Термин **«гетерохроматин»** используют для обозначения хромосомных областей, обширных настолько, что их можно наблюдать под микроскопом как формацию с особо компактной структурой. Оба типа изменений (умолкание или переход в гетерохроматин), по сути, одно и то же. Они обусловлены тем, что с хроматином связываются дополнительные белки и, прямо или косвенно, мешают факторам транскрипции и РНК-полимеразе активировать промоторы в данной области.

Инициируется или нет транскрипция конкретного гена, зависит от изменений (активирующих либо репрессирующих) состояния его промотора. Эти изменения зависят от взаимодействия гистонов с другими белками и с ДНК. Изменения в структуре хроматина инициируются модификацией N-концевых «хвостов» в молекулах гистонов, в особенности H3 и H4. «Хвосты» состоят из 15–30 аминокислот N-концевой области каждого из четырех гистонов сердцевины; у гистона H2A такой «хвост» есть также на C-конце. «Хвосты» H2B и H3 пропущены между сверхспиральными витками нуклеосомной ДНК (см. рис. 29.21 в разд. 29.7 «Организация гистонового октамера). На РИС. 30.8 показано, как эти «хвосты» могут быть модифицированы по нескольким сайтам посредством метилирования, ацетилирования или фосфорилирования. Другие модификации, такие как убиквитинилирование или пришивания полипептида sumo, тоже возможны, но не так хорошо изучены.

Ацетилирование и метилирование идут по свободной (ε) аминогруппе лизина. Как видно из РИС. 30.9, ацетилирование нейтрализует положительный заряд,



РИС. 30.8 N-концевые «хвосты» гистонов H3 и H4 могут быть ацетилированы, метилированы или фосфорилированы по нескольким положениям

присутствующий на радикале лизина в NH_3^+-форме. Метилироваться кроме лизина может аргинин. Фосфорилирование гистонов происходит по гидроксильной группе серина или треонина; добавление фосфатной группы приносит отрицательный заряд. Лизин может быть моно-, ди- или триметилирован (во всех трех случаях положительный заряд на радикале лизина сохраняется; аргинин может быть моно- или диметилирован (симметрично или асимметрично).

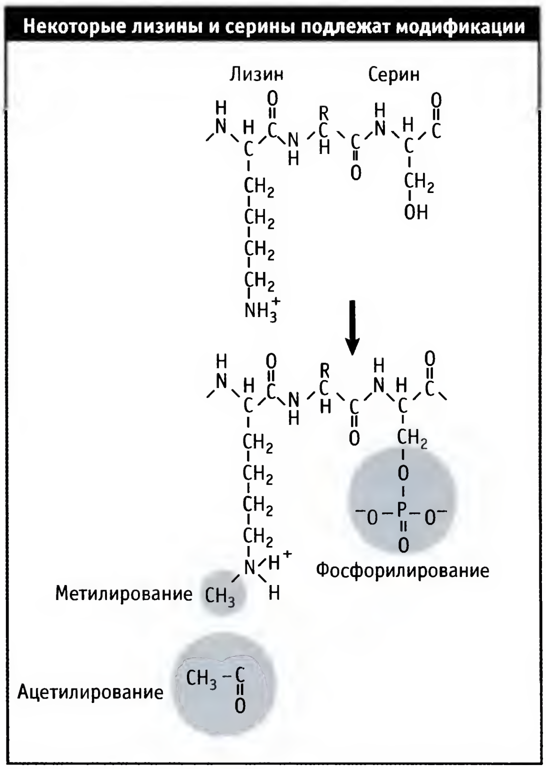


РИС. 30.9 Ацетилирование лизина или фосфорилирование серина снижает общий положительный заряд белка

Перечисленные модификации гистонов носят временный характер. Они могут изменить заряд молекулы белка и, следовательно, создают потенциал изменения свойств нуклеосомных октамеров. Модификация гистонов ассоциирована со структурными изменениями, происходящими в хроматине при репликации или транскрипции. Фосфорилирование разных гистонов по специфическим положениям необходимо для осуществления конкретных процессов. Например, положение ^{10}Ser в гистоне H3 фосфорилируется при конденсации хромосом во время митоза.

В синхронизированных по клеточному циклу клетках культуры и предсуществующие, и новосинтезированные гистоны сердцевины в S-фазе, по всей видимости, ацетилированы и метилированы. Далее по ходу клеточного цикла гистоны теряют эти ацетильные и метильные группы. Совпадение модификации гистонов во времени с репликацией говорит о том, что ацетилирование и метилирование гистонов могут иметь отношение к сборке нуклеосом. Было сделано предположение, что уменьшение положительных зарядов на гистонах может снижать их сродство к ДНК, способствуя более тщательному контролю сборки нуклеосом. Однако это предположение оказалось не очень состоятельно ввиду того, что нуклеосомы могут быть смонтированы (во всяком случае, *in vitro*) из немодифицированных гистонов. Ацетилирование гистонов принципиально важно для сборки нуклеосом у дрожжей и, вероятнее всего, требуется для некоторых белок-белковых взаимодействий, происходящих на поздних стадиях сборки (см. разд. 30.6 «Ацетилирование гистонов происходит в двух случаях»).

Фосфорилирование и дефосфорилирование гистона H1 происходят в режиме, отличном от модификационного цикла других гистонов. В культивируемых клетках млекопитающих, в S-фазе, в H1 вносится по одной или по две фосфатных группы на молекулу. Позднее, при митозе, происходит более интенсивное фосфорилирование: число фосфатов в молекуле H1 доводится до шести. Под конец клеточного деления происходит удаление из H1 всех фосфатных групп. Фосфорилирование H1 катализирует киназа M-фазы, играющая в клеточном цикле роль принципиального переключателя на митоз. На самом деле, этот фермент часто исследуют на предмет его H1-киназной активности. О фосфатазах, удаляющих фосфаты с H1 в конце митоза, известно немного.

Момент добавления к гистону H1 большинства его фосфатов породил догадки о будто бы причастности фосфорилирования H1 к митотической конденсации хромосом. Для одноклеточного простейшего *Tetrahymena* эти догадки не получили подтверждения, поскольку делеция гена *H1* нарушает общие свойства хроматина лишь незначительно. В частности, способность хроматина конденсироваться при митозе если и страдает от отсутствия H1, то довольно слабо. В целом при делеции *H1* некоторые гены активируются, другие же оказываются репрессированными, что говорит о разнообразных локальных изменениях структуры хроматина. Мутации, элиминирующие в H1 сайты фосфорилирования, фенотипа не имеют, зато мутации, имитирующие фосфорилирование, по своему фенотипу похожи на делецию *H1*. Это говорит о том, что фосфорилирование H1 нейтрализует влияние этого гистона на локальную структуру хроматина.

Искажают ли модификации гистонов непосредственно структуру нуклеосом, или эффект скорее косвенный и адресован более высоким уровням организации хроматина? Данных, которые указывали бы на разницу в свойствах нуклеосом, обусловленную именно модификацией гистонов, существует немного. Между тем, описано уже несколько случаев, когда благодаря модификации у гистонов возникают сайты связывания с негистоновыми белками, что меняет свойства хроматина.

Выборка нуклеосом, в которых гистоны подлежат модификации, может варьировать. Модификация гистонов может быть локальным событием, ограничивающимся, например, нуклеосомами одного промотора, и может носить характер глобального мероприятия в масштабах, например, целой хромосомы. На РИС. 30.10 показана общая закономерность, в соответствии с которой ацетилирование гистонов ассоциировано с активным хроматином, метилирование гистонов ассоциировано с неактивным хроматином. Это, однако, не такое простое правило; важны конкретные сайты, которые модифицированы (точно так же, как комбинации сайт-специфичных модификаций), и, конечно же, из общей закономерности есть исключения. Так, гистоны, метилированные по конкретному положению, обнаруживаются в активном хроматине. Мутации в одном из дрожжевых гистонацетилазных комплексов, против ожидания, препятствуют сайленсингу некоторых генов, что подчеркивает отсутствие единства среди эффектов ацетилирования.

Специфичность модификаций усугубляется тем, что многие из ферментов, катализирующих модификации, имеют свои индивидуальные сайты-мишени



РИС. 30.10 Ацетилированное состояние «хвостов» H3 и H4 характерно для активного хроматина, а метилированное — для неактивного хроматина

Модификация гистонов влияет на структуру и функцию хроматина

Гистон	Сайт	Модификация	Функция
Н3	Lys-4	Метилирование	Активация
Н3	Lys-9	Метилирование	Конденсация хроматина
	Lys-9	Метилирование	Требуется для метилирования ДНК
	Lys-9	Ацетилирование	Активация
Н3	Ser-10	Фосфорилирование	Активация
Н3	Lys-14	Ацетилирование	Препятствует метилированию по Lys-9
Н3	Lys-79	Метилирование	Сайленсинг в теломерах
Н4	Arg-3	Метилирование	Неизвестно
Н4	Lys-5	Ацетилирование	Сборка
Н4	Lys-12	Ацетилирование	Сборка
Н4	Lys-16	Ацетилирование	Сборка нуклеосом
	Lys-16	Ацетилирование	Активация Х-хромосомы у мух

РИС. 30.11 Большинство сайтов модификации в «хвостах» гистонов подлежит модификации единственного типа. Многие модификации гистонов имеют регуляторный смысл

на конкретных гистонах. В таблице на РИС. 30.11 перечислены эффекты некоторых модификаций. Большинство модифицированных сайтов не подлежат альтернативным модификациям. В некоторых случаях модификация одного сайта может активировать или ингибировать модификацию другого сайта. Комбинации модификаций-сигналов, которые, согласно одной из концепций, могут использоваться в клетке для обозначения типов хроматина, иногда называют *гистоновым кодом*.

30.6 Ацетилирование гистонов происходит в двух случаях

Ключевые понятия:

- Кратковременное ацетилирование гистонов происходит в процессе репликации.
- Вне репликации, ацетилирование гистонов ассоциировано с активацией экспрессии генов.

Все гистоны нуклеосомной сердцевины могут быть ацетилированы. Главные мишени для ацетилирования — это лизины в N-концевых «хвостах» гистонов Н3 и Н4. Ацетилирование происходит при двух разных стечениях обстоятельств:

- во время репликации ДНК или
- когда гены активируются.

Когда в S-фазе клеточного цикла геномная ДНК реплицируется, гистоны временно ацетилированы. Как показано на РИС. 30.12, ацетилирование происходит до включения гистонов в нуклеосомы. Известно, что гистоны Н3 и Н4 к тому времени, когда они ассоциируют и образуют тетрамер $H3_2$–$H4_2$, уже ацетилированы. Тетрамер включается в состав нуклеосомы и вскоре после этого теряет свои ацетильные группы.

О важности ацетилирования гистонов при репликации говорит тот факт, что если одновременно предотвратить ацетилирование Н3 и Н4, то дрожжевые клетки теряют жизнеспособность. В качестве субстратов для ацетилирования Н3 и Н4 дублируют друг друга, поскольку дрожжи отлично обходятся возможностью ацетилировать во время S-фазы, хотя бы один из этих гистонов. Ацетилированию отводят две возможных роли: возмож-

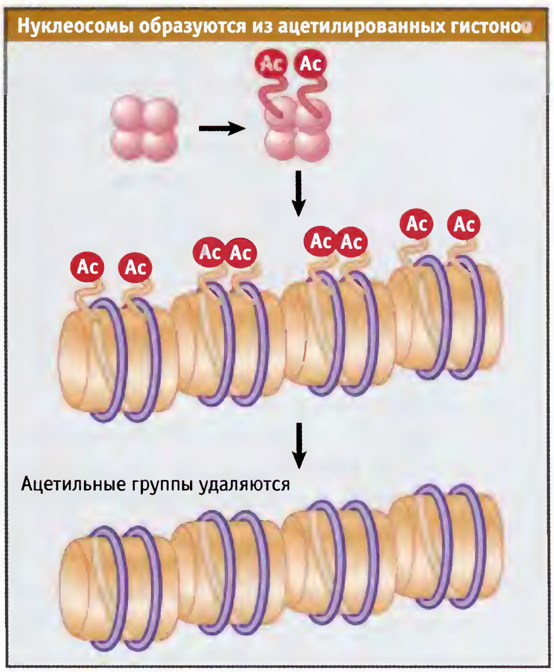


РИС. 30.12 Во время репликации ацетилирование гистонов происходит до их включения в нуклеосомы

но, оно требуется для узнавания гистонов факторами сборки нуклеосом, либо оно необходимо для самого процесса сборки и/или придания нуклеосомам особенной структуры.

Факторы, про которые известно, что они участвуют в сборке хроматина (нуклеосом), не отличают ацетилированнные гистоны от неацетилированных и используют их без разбора. Следовательно, скорее всего, наличие ацетильной группы требуется для взаимодействий на последующих этапах сборки. Долгое время считалось, что ацетилирование необходимо для контроля белок-белковых взаимодействий, возникающих по мере того, как гистоны включаются в состав нуклеосом. В пользу именно этой роли говорит поведение гистонацетилазы SAS у дрожжей. Комплекс SAS связывает комплексы сборки хроматина в репликативной вилке, где он ацетилирует ^{16}Lys гистона H4. Возможно, работа комплекса SAS — это часть клеточной «программы» по учреждению паттернов ацетилирования гистонов после репликации.

Вне репликации и S-фазы, статус ацетилирования гистонов хроматина в целом коррелирует со статусом экспрессии генов, которые в нем содержатся. В хромосомном домене, содержащем активные гены, частота ацетилирования гистонов повышена; в то же время хроматин, ацетилированный по гистонам, более чувствителен к ДНКазе I, а зачастую и к микрококковой нуклеазе. На РИС. 30.13 показано, что мишенью для ацетилирования служат «хвосты» гистонов в нуклеосомах. Этим отчасти объясняется, почему активный хроматин более чувствителен к нуклеазам: во многом из-за ацетилированного состояния нуклеосом в области промотора и его окрестностей, когда ген активирован.

Кроме событий, носящих локальный характер и относящихся к отдельным промоторам, изменения статуса ацетилирования бывают широкомасштабными. Это особенно характерно для половых хромосом и является частью механизма, посредством которого происходит коррекция избыточной активности генов, вызванной присутствием двух X-хромосом в клетках особей одного пола, по сравнению с единственной X-хромосомой, в комплекте с Y-хромосомой, в клетках особей другого пола (см. разд. 31.5 «X-хромосомы претерпевают глобальные изменения»). В неактивной X-хромосоме самок млекопитающих у гистона H4 степень ацетилирования понижена. Напротив, в «суперактивной» X-хромосоме у самцов *Drosophila* степень ацетилирования гистона H4 повышена. Это говорит о том, что присутствие ацетильных групп на гистонах нуклеосом может служить предпосылкой более рыхлой (менее конденсированной), «активной», структуры. У самцов *Drosophila* X-хромосома ацетилирована прицельно по ^{16}Lys гистона H4. За это отвечает гистонацетилаза (гистонацетилтрансфераза) MOF; подход к хромосоме MOF совершает в составе большого белкового комплекса. Этот комплекс отвечает за внесение в X-хромосому общих изменений, придающих ее генетическому материалу способность экспрессироваться с повышенной интенсивностью. Массированное ацетилирование гистона H4 — это лишь одна из активностей этого комплекса.

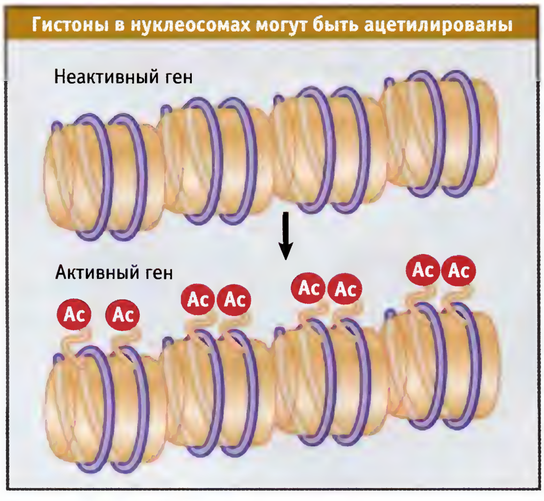


РИС. 30.13 Ацетилирование, ассоциированное с активацией генов, происходит с гистонами непосредственно в нуклеосомах

30.7 Гистонацетилазы ассоциированы с активаторами транскрипции

Ключевые понятия:

- Деацетилированному хроматину свойственна более рыхлая (деконденсированная) структура.
- Активаторы транскрипции ассоциированы с гистонацетилазами в составе больших белковых комплексов.
- Гистонацетилазы различаются по специфичности к субстрату (мишени).
- Ацетилирование гистонов может влиять на транскрипцию количественно или качественно.

Ацетилирование гистонов обратимо. Каждое из направлений реакции катализируется особым ферментом. Ферменты, способные ацетилировать гистоны, называют гистонацетилазами, или (более правильно) **гистонацетилтрансферазами** (histone-acetyltransferases, HATs). Удалить у гистона ацетильную группу могут другие ферменты, **гистондеацетилазы** (histone-deacetylases, HDACs). Гистонацетилтрансферазы можно подразделить на две группы. Ферменты группы A действуют на гистоны, уже вошедшие в состав хроматина, и участвуют в контроле транскрипции. Ферменты группы B работают с новосинтезированными гистонами в цитоплазме и имеют отношение сборке нуклеосом.

Для изучения обратимого ацетилирования гистонов оказались весьма полезными два вещества: трихостатин и масляная кислота. Эти вещества ингибируют

гистондеацетилазы; с присутствием любого из них связано накопление в клетке ацетилированных нуклеосом. В экспериментах, где использовались эти ингибиторы, с успехом прошел проверку общий тезис о взаимосвязи ацетилирования гистонов с экспрессией генов. На самом деле, способность масляной кислоты вызывать в хроматине изменения, подобные тем, какие бывают при активации генов, как раз и послужили первыми указаниями на взаимосвязь ацетилирования с активностью генов.

Большой прорыв в понимании смысла ацетилирования гистонов был связан с идентификацией ацетилирующих и деацетилирующих ферментов, а также с выявлением взаимоотношений этих ферментов с другими белками-участниками событий активации и репрессии, обладающими специфичностью к конкретным генам. Коренное изменение в видении смысла ацетилирования гистонов было вызвано открытием, что гистонацетилазы — это не специальные ферменты, постоянно ассоциированные с хроматином. Как выяснилось, гистонацетилтрансферазная активность свойственна многим известным активаторам транскрипции.

Зацепка появилась, когда каталитическая субъединица гистонацетилтрансферазы группы А была идентифицирована как гомолог регуляторного белка Gcn5 дрожжей. Последующие эксперименты показали, что гистонацетилтрансферазной активностью обладает сам Gcn5, с гистонами Н3 и Н4 в качестве субстратов. Белок Gcn5 является частью комплекса-адаптора, необходимого для взаимодействия некоторых энхансеров с промоторами их генов-мишеней. В частности, для активации гена-мишени необходима присущая комплексу-адаптору активность гистонацетилтрансферазы.

С учетом этих новых знаний можно обновить всю картину работы активаторов транскрипции так, как это показано на РИС. 30.14. РНК-полимераза связана с гиперчувствительным сайтом, а коактиваторы ацетилируют гистоны окрестных нуклеосом. Конкретных примеров такого рода на сегодняшний день известно довольно много.

Нить расследования функции Gcn5 приводит к одному из самых важных гистонацетилазных комплексов. У дрожжей Gcn5 является частью 1,8 МДа белкового комплекса. Этот Spt-Ada-Gcn5 ацетилтрансферазный комплекс, сокращенно SAGA, содержит несколько белков, участвующих в транскрипции, в том числе несколько факторов TAF_{II}. Гистонацетилтрансферазой также является субъединица фактора $TF_{II}D$ дрожжей, $TAF_{II}145$. (Вместе с гомологом, $TAF_{II}250$ млекопитающих, они известны под именем TAF1.) Функции $TF_{II}D$ и SAGA отчасти перекрываются; примечательно, что дрожжи вполне способны обходиться без Gcn5 либо $TF_{II}D$, но страдают от делеции обоих. Это говорит о том, что для нормальной экспрессии генов гистонацетилтрансферазная активность значит очень много, и обеспечить ее в достаточном объеме готов в одиночку любой из комплексов, SAGA или $TF_{II}D$. Как можно было ожидать, судя по размерам комплекса SAGA, ацетилирование гистонов — только одна из его функций, хотя другие роли комплекса SAGA в активации генов изучены не так хорошо.

Одним из первых активаторов транскрипции, охарактеризованных в качестве гистонацетилтрансфераз, был p300/CBP (где CBP — это сокращение от CREB-binding protein). Вообще-то, p300 и CBP — это разные белки, но они так похожи, что их часто объединяют. p300/CBP работает коактиватором, то есть наводит связь между акиватором и аппаратом транскрипции (см. рис. 25.7). В этом качестве p300/CBP работает с различными активаторами, в числе которых рецепторы гормонов, AP-1 (c-Jun и c-Fos), MyoD. Коактивация может быть ингибирована регуляторными белками вирусного происхождения, E1A аденовируса или Т-антигеном SV40, которые связываются с p300/CBP и препятствуют его взаимодействию с факторами транскрипции; в этом, собственно, и заключается механизм подавления клеточной транскрипции указанными вирусами. Ингибирование функции p300/CBP — важная составляющая вклада этих вирусов в доведение клетки до опухолеродного состояния. p300/CBP ацетилирует N-концевые «хвосты» молекул гистона H4 в составе нуклеосом. Другой коактиватор, PCAF, предпочитает ацетилировать в нуклеосомах молекулы H3. p300/CBP и PCAF образуют комплекс, задействованный в активации транскрипции. Другая гистонацетилтрансфераза — коактиватор ACTR,

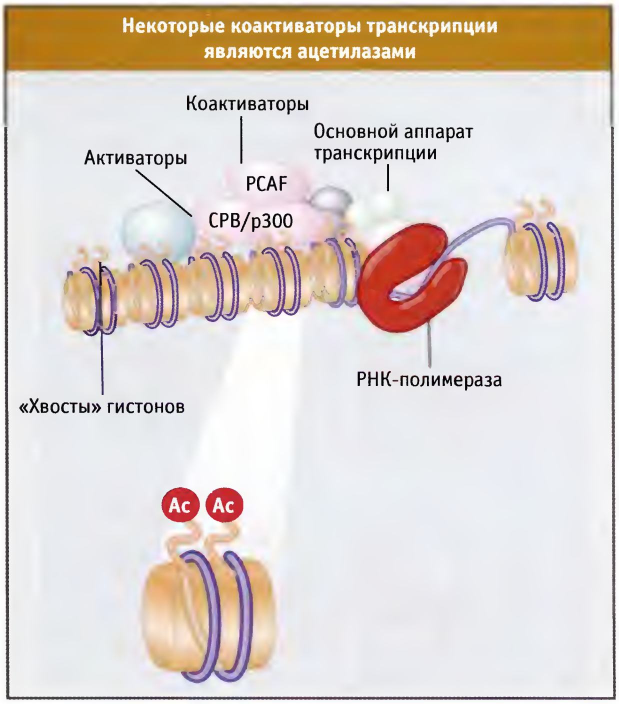


РИС. 30.14 Коактиватор может обладать гистонацетилазной активностью и ацетилировать «хвосты» нуклеосомных гистонов

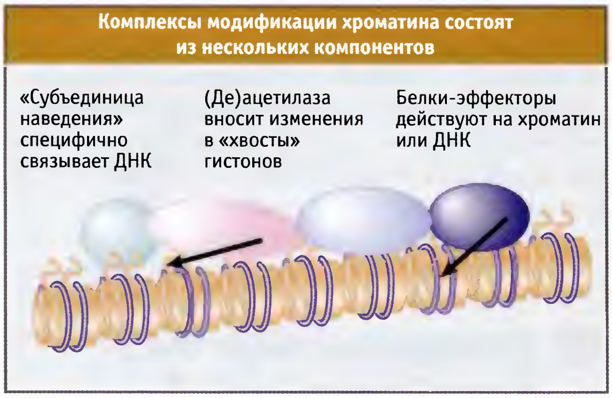


РИС. 15 **Белковые комплексы, модифицирующие структуру или активность хроматина, имеют несколько субъединиц: для наведения на конкретное место действия, для ацетилирования или деацетилирования гистонов, для других воздействий на хроматин или ДНК**

работающий с рецепторами гормонов и не только сам ацетилирующий гистоны Н3 и Н4, но также привлекающий p300/CBP и PCAF к вступлению в коактивирующий комплекс. Одновременное присутствие множества гистонацетилазных активностей в составе коактивирующего комплекса можно объяснить различиями в их специфичности к субстрату и потребность во множестве разных событий ацетилирования, для того, чтобы «разбудить» экспрессию.

Гистонацетилтрансферазы группы А, что характерно, функционируют в составе больших белковых комплексов. На РИС. 30.15 представлена упрощенная модель «поведения» такого комплекса. Доставка ацетилтрансферазных комплексов по месту назначения на ДНК может координироваться посредством взаимодействия комплексов с ДНК-связывающими факторами: они диктуют гистонацетилтрансферазе выбор мишени. В комплексе есть и субъединицы-эффекторы; их задача — вмешательство в структуру хроматина или непосредственно в транскрипцию. Похоже, что эффекторы (по крайней мере, некоторые из них) нуждаются в том, чтобы в поле их деятельности гистоны были ацетилированы. Соответственно, огромную роль в регуляции экспрессии генов играет деацетилирование гистонов, катализируемое гистондеацетилазами.

Ацетилирование гистонов имеет место как при репликации (на короткий промежуток времени), так и при транскрипции (в этом случае оно поддерживается вплоть до завершения геном своей активности). Играет ли ацетилирование в каждом из этих случаев одну и ту же роль? Одна из возможностей состоит в том, что эффект ацетилирования гистонов, универсальный и однотипный, выражается через структуру нуклеосом. Ацетилирование может работать как «разрыхлитель» нуклеосомной сердцевины. Соответственно, при репликации ацетилированные гистоны могут быть с большей легкостью вмонтированы в новые формирующиеся сердцевины, а при транскрипции эффект того же рода может быть применен с результатом вплоть до смещения гистонового октамера с ДНК. Даже если смещения не предполагается, ацетилирование гистонов может подготавливать сайты ДНК к связыванию белков, необходимые для транскрипции. В каждом случае эффект ацетилирования обращается вспять деацетилированием.

Эффект ацетилирования — количественный или качественный? В одних ситуациях, для срабатывания некоторого эффекта некоторой совокупности гистонов необходимо набрать некоторое заданное число ацетильных групп, и какие именно позиции займут эти ацетильные группы, по большому счету, не существенно. В ситуациях противоположного типа, каждое конкретное событие ацетилирования имеет эксклюзивную значимость. Существование комплексов с множественной гистонацетилтрансферазной активностью можно интерпретировать с учетом ситуаций обоих типов: если ферменты имеют разную специфичность, то для того, чтобы ацетилировать гистоны по достаточному числу положений и достаточно быстро, нужно использовать все доступные активности; или же активность каждого отдельно взятого фермента имеет для транскрипции свой особый смысл. Похоже, что при репликации уместна, скорее, количественная модель (по крайней мере, для гистона Н4) ацетилирование по любым двум позициям (из трех доступных) функционально адекватно (то есть придает молекуле Н4 годность к инкорпорации в нуклеосому). В другом клеточном контексте, когда изменения структуры хроматина сопутствуют изменению статуса экспрессии области генома, и ацетилирование гистонов призвано оказать влияние на транскрипцию, очень важно, по каким именно положениям в молекулах гистонов идет это ацетилирование (см. разд. 31.3 «Существование гетерохроматина обусловлено взаимодействия гистонов с другими белками»).

30.8 Гистондеацетилазы ассоциированы с репрессорами транскрипции

Ключевые понятия:

- Деацетилирование гистонов ассоциировано с репрессией активности генов.
- Гистондеацетилазы состоят в комплексах с белками-репрессорами.

У дрожжей мутации по *SIN3* и *RPD3* ведут себя так, как будто продукты этих локусов репрессируют множество генов. В действительности, их белки вступают в комплекс с ДНК-связывающим белком Ume6, который, в свою очередь, связывается с элементом *URS1*. Комплекс репрессирует транскрипцию на промоторах, содержащих *URS1*, как это показано на РИС. 30.16. Rpd3 «умеет» деацетилировать гистоны; насчет Sin3 пока не

понятно: то ли он должен лишь доставлять Rpd3 на промотор, то ли он играет в репрессии какую-то дополнительную роль.

Похожая система для репрессии транскрипции была обнаружена в клетках млекопитающих. К семейству регуляторов транскрипции bHLH относятся белки-активаторы, функционирующие в форме гетеродимеров. Типичный представитель семейства — MyoD (см. разд. 25.15 «Белки с доменами «спираль-петля-спираль» димеризуются в различных сочетаниях»). К тому же семейству bHLH относятся и некоторые репрессоры, в частности, гетеродимер Mad:Max, (где Mad — это любой представитель группы близкородственных белков Mad). Гетеродимер Mad:Max связывается с конкретными сайтами ДНК и взаимодействует с гомологом Sin3 (который у мыши называется mSin3, а у человека hSin3). Белок mSin3 входит в состав комплекса-репрессора, компонентами которого также являются гистонсвязывающие белки и гистондеацетилазы, HDAC1 и HDAC2. Активность деацетилаз необходима для репрессии. Функциональная самостоятельность (модульность) этой системы подчеркивается другими вариантами ее использования: корепрессор SMRT, помогающий рецепторам ретиноидных гормонов репрессировать некоторые гены-мишени, выполняет свою функцию, связываясь с mSin3, который, в свою очередь, рекрутирует HDAC с их гистондеацетилазными активностями. Другим способом привлечь HDAC к работе на конкретном сайте может быть контакт с MeCP2, белком, который связывается с метилированными цитозинами (см. разд. 24.19 «CpG-островки являются мишенями для регуляции»).

Отсутствие ацетилирования по гистонам является также отличительной чертой гетерохроматина. Это верно и для конститутивного гетерохроматина (к которому, как правило, относятся области центромер и теломер), и для факультативного гетерохроматина (то есть областей, специфически инактивированных в одной клетке, но способных при этом быть активными в какой-нибудь другой клетке). Как правило, в областях гетерохроматина N-концевые «хвосты» гистонов H3 и H4 не ацетилированы.

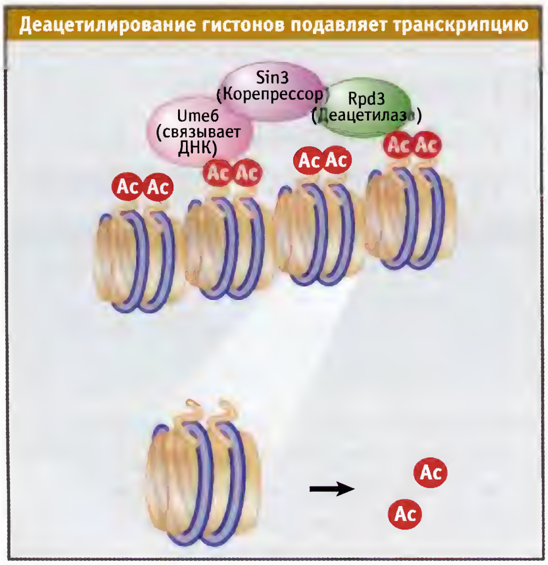


Рис. 30.16 В комплекс репрессии входят три компонента: ДНК-связывающая субъединица, корепрессор и гистондеацетилаза

30.9 Метилирование гистонов и метилирование ДНК взаимосвязаны

Ключевые понятия:

- Метилированность — и по гистонам, и по ДНК — отличительная черта неактивного хроматина.
- События метилирования этих двух субстратов могут быть взаимосвязаны.

В хроматине, где метилированы и гистоны, и ДНК, гены, как правило, не экспрессируются. Метилирование гистонов происходит по конкретным аминокислотным положениям: в частности, по двум лизинам в «хвосте» H3 и аргинину в «хвосте» H4.

Метилирование ^{9}Lys в гистоне H3 является отличительной чертой конденсированных областей хроматина (в частности, гетерохроматина, в общей массе) и менее крупных областях, про которые тоже известно, что они не экспрессируются. Фермент гистонметилтрансферазу, мишенью которому служит ^{9}Lys, называют SUV39H1 (почему его так причудливо назвали, мы узнаем из разд. 30.13 «У белков, модифицирующих хроматин, есть общие аминокислотные мотивы»). Активный центр этого фермента содержит область, известную как SET-домен. Другим гистонметилтрансферазам мишенью служит аргинин. Вдобавок, метилирование может идти по ^{79}Lys в глобулярной сердцевине молекулы H3; эта реакция может быть необходима для формирования гетерохроматина теломер.

До недавних пор считалось, что метилирование гистонов необратимо. На сегодня, однако, было идентифицировано уже несколько гистондеметилаз; в их числе — лизин-специфичная деметилаза LSD1, действующая на ^{4}Lys гистона H3, а также фермент, который деметилирует аргинин в гистонах H3 и H4. Как регулируется деметилирование, пока неясно.

Что касается ДНК, большинство сайтов метилирования в ней — это CpG-островки (см. разд. 24.19 «CpG-островки являются мишенями для регуляции»). В гетерохроматине последовательности CpG, как правило, метилированы. Наоборот, CpG-островкам, находящимся в активных промоторах, свойственно *не* быть метилированными (см. разд. 24.18 «Экспрессия генов связана с деметилированием»).

Процессы метилирования ДНК и метилирования гистонов взаимодействуют по типу положительной

обратной связи. Метилирование гистона H3 служит сигналом, привлекающим ДНК-метилазу в хроматин. Порядок событий таков, что сначала ^{9}Lys подвергается деацетилированию и тем самым превращается в субстрат для метилирования. Затем H3 переводится в состояние Me^{9}Lys или Me$_2$^{9}Lys, обеспечивая ДНК-метилазе сайт для связывания. Некоторые гистонметилтрансферазы имеют потенциальные сайты для связывания метилированного дублета CpG: возможно, метилированная ДНК производит свой усиливающий эффект (в рамках контура положительной обратной связи с метилированием гистонов) за счет того, что обеспечивает субстрат гистонметилтрансферазам. Таким образом, метилирование одного субстрата (ДНК или гистонов) срабатывает на запуск метилирования другого. Это взаимодействие на уровне процессов — очень распространенное, поскольку имеются доказательства его существования в клетках грибов, растений и животныхх— для регуляции транскрипции на промоторах, используемых РНК-полимеразами I и II, а также для поддержания гетерохроматина в его инертном состоянии.

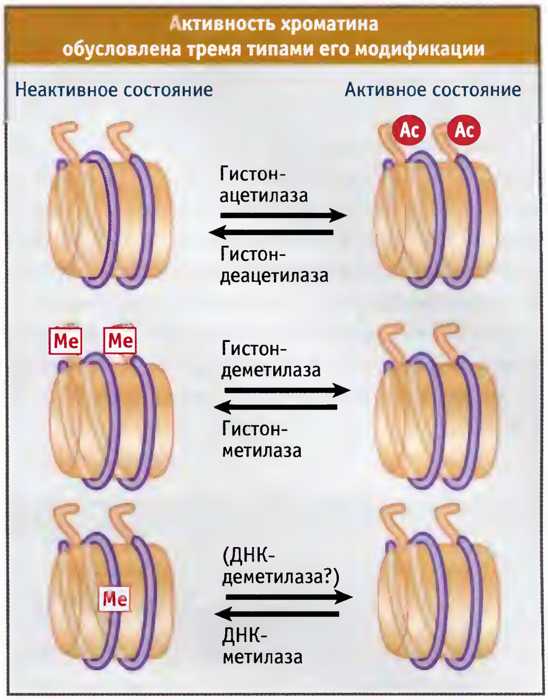


РИС. 30.17 Ацетилирование гистонов активирует хроматин, метилирование ДНК и гистонов инактивирует хроматин

30.10 Переходы между состояниями хроматина происходят через модификацию его компонентов

Ключевые понятия:

- **Ацетилирование гистонов связано с активацией генов.**
- **Метилирование ДНК и/или гистонов характерно для неактивного хроматина.**

На РИС. 30.17 просуммированы три типа различий, обнаруживаемых между активным и неактивным хроматином:

- Активный хроматин ацетилирован по «хвостам» гистонов H3 и H4.
- Неактивный хроматин метилирован по ^{9}Lys гистона H3.
- Неактивный хроматин метилирован по цитозинам дублетов CpG.

Обратный ход событий можно пронаблюдать, сравнив процессы активации промотора и возникновения гетерохроматина. Действия ферментов, модифицирующих хроматин в каждом из этих случаев, вселяют уверенность в том, что события активации и события инактивации строго альтернативны. Так, метилирование гистона H3 по ^{9}Lys и его же ацетилирование по ^{14}Lys не могут идти одновременно.

Ацетилазы и деацетилазы могут запускать инициирующие события. Деацетилирование позволяет произойти метилированию, что вызывает образование гетерохроматического комплекса (см. разд. 31.3 «Существование гетерохроматина обусловлено взаимодействиями гистонов с другими белками»). Ацетилирование гистонов помечает область хроматина как активную (см. разд. 30.11 «Активация промотора представляет собой упорядоченную серию событий»).

30.11 Активация промотора представляет собой упорядоченную серию событий

Ключевые понятия:

- **Комплекс ремоделирования хроматина может рекрутировать комплекс ацетилирования гистонов.**
- **Ацетилирование гистонов стабилизирует хроматин в активированном состоянии.**

Каким же образом гистонацетилазы и гистондеацетилазы оказываются привлеченными к конкретным мишеням? Как было показано для ремоделирующего комплекса, процесс этот опосредованный. Активатор (или репрессор), специфичный по отношению к нуклеотидной последовательности, может, через взаимодействие с компонентом ацетилазного (или деацетилазного) комплекса, привлечь его к промотору.

Кроме того, комплексы ремоделирования хроматина и комплексы модификации гистонов могут взаимодействовать между собой напрямую. Привлечение к промотору ремоделирующего комплекса SWI/SNF может привести к приходу на этот промотор ацетилазного комплекса SAGA. Последующее ацетилирование гистонов может стабилизировать ассоциацию промотора с комплексом SWI/SNF, создавая взаимно-подкрепляющие отношения между действиями компонентов белкового аппарата инициации транскрипции.

Все события, которые имеют место на промоторе, можно выстроить в цепочку, представленную на РИС. 30.18. Инициирующим событием выступает связывание компонента, специфичного по отношению к нуклеотидной последовательности сайта ДНК (и способного разыскать свою мишень в контексте хроматина). Он рекрутирует комплекс ремоделирования; в нуклеосомной организации промотора происходят изменения. Эти изменения позволяют связаться ацетилирующему комплексу, и уже ацетилированность гистонов-мишеней в дальнейшем служит ковалентной маркировкой активированного локуса.

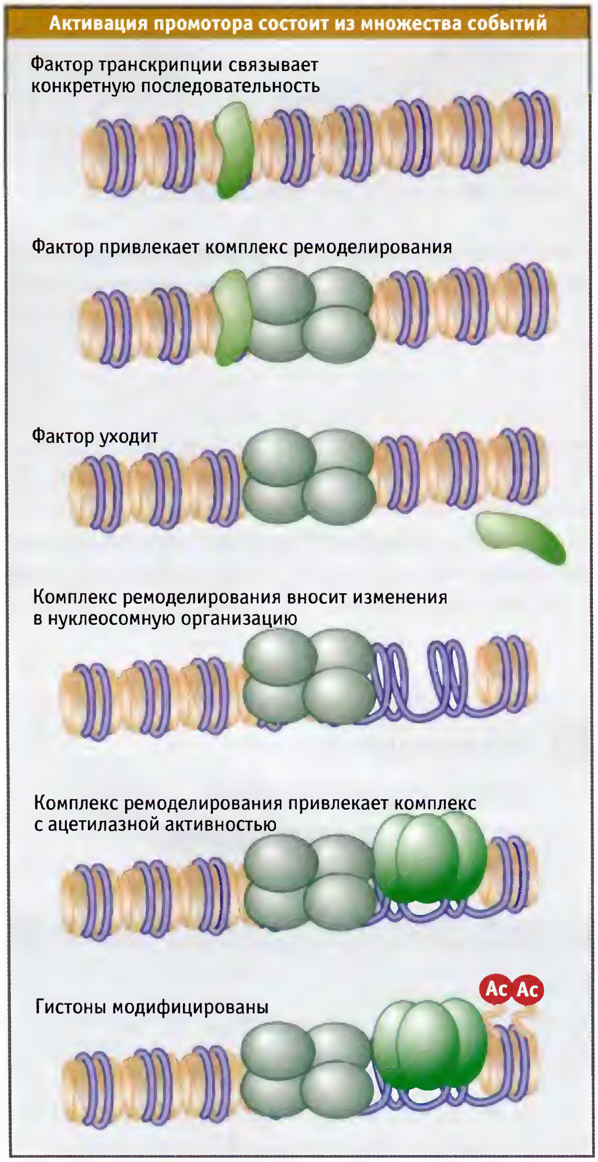


РИС. 30.18 Активация промотора начинается с узнавания его нуклеотидной последовательности специфичным белком-активатором. Затем к работе на промоторе привлекаются комплексы ремоделирования и ацетилирования. Точный порядок событий для разных промоторов может отличаться

Также в промоторе происходит модификация ДНК. Метилированность цитозинов в дублетах CpG означает отсутствие экспрессии (см. разд. 24.18 «Экспрессия генов связана с деметилированием»). Основания для опознания участка ДНК в качестве мишени для метилирования известны лишь приблизительно (см. разд. 31.8 «Метилирование ДНК обусловливает импринтинг»).

Понятно, что ремоделирование хроматина в области промотора требует разнообразных модификаций нуклеосом, в том числе ацетилирования гистонов; а какие изменения должны произойти в самом гене, чтобы РНК-полимераза смогла его прочитать? Известно, что РНК-полимераза может транскрибировать ДНК *in vitro* со скоростью, сопоставимой с реальными измерениями скорости транскрипции *in vivo* (25 нуклеотидов в секунду), только на матрице свободной (*не* организованной в виде нуклеосом) ДНК. Несколько белков отличаются своей способностью увеличивать скорость прочтения ДНК РНК-полимеразой. Их общей особенностью является то, что естественным местом их работы является хроматин. Модель работы этих белков, принятая на сегодня, заключается в том, что они вступают в контакт с РНК-полимеразой и переносятся вместе с ней вдоль матрицы, попутно внося изменения в нуклеосомную организацию гена через воздействие на гистоны. Среди катализирующих белков есть и гистонацетилазы. По одной из версий, первая РНК-полимераза, прочитывающая ген, выступает в роли первопроходца, неся с собой факторы, меняющие структуру единицы транскрипции, чтобы следующим полимеразам пройти было уже легче.

30.12 Фосфорилирование гистонов влияет на структуру хроматина

Ключевые понятия:

- **Из четырех гистонов нуклеосомной сердцевины мишенями для фосфорилирования служат как минимум два.**

Фосфорилирование гистонов имеет место при одном из двух обстоятельств:

- в ходе клеточного цикла, или
- в связи с ремоделированием хроматина.

Давно известно, что гистон H1 фосфорилируется при митозе; однако сравнительно недавно было открыто,

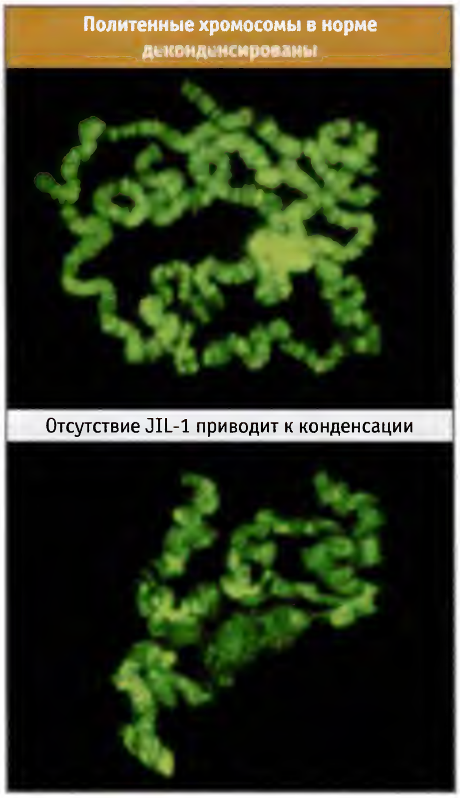


РИС. 30.19 У мух, не имеющих киназы JIL-1, политенные хромосомы выглядят уродливо: они слишком конденсированы. Фотография любезно предоставлена Jorgen Johansen and Kristen M. Johansen, Iowa State University

что H1 чрезвычайно хорош как субстрат для киназы Cdc2, осуществляющей контроль над клеточным делением. Это открытие спровоцировало тезис о том, что фосфорилирование H1 может быть взаимосвязано с конденсацией хроматина. Однако прямых доказательств того, что Cdc2 фосфорилирует H1 *in vivo*, получено не было, так что неясно, играет оно какую-нибудь роль в делении клетки, или нет.

Утрата киназы, фосфорилирующей гистон H3 по ^{10}Ser, губительна для структуры хроматина. На РИС. 30.19 проведено сравнение обычной, расправленной, структуры набора политенных хромосом *Drosophila melanogaster* (верхняя фотография) с аналогичной структурой из клеток мутанта, у которого отсутствует функция киназы JIL-1 (нижняя фотография). Отсутствие JIL-1 летально, но хромосомы можно успеть визуализировать до того, как личинки-мутанты погибнут.

Нарушение структуры хроматина в этом случае вызвано, скорее всего, невозможностью фосфорилировать гистон H3 (хотя, конечно же, у JIL-1 могут быть и другие мишени). Это указывает на необходимость фосфорилирования H3 для придания хроматину более расправленной структуры, свойственной областям эухроматина. Экспериментальное подтверждение тому, что JIL-1 действует непосредственно на хроматин, состоит в том, что эта киназа ассоциирована с белковым комплексом, который связывается с X-хромосомой, чтобы интенсифицировать экспрессию ее генов у самцов *Drosophila* (см. разд. 31.5 «X-хромосомы претерпевают глобальные изменения»).

Впечатление о роли фосфорилирования гистонов, таким образом, довольно противоречивое. Если фосфорилирование гистонов — важная составляющая клеточного цикла, то, скорее всего, оно функционирует как сигнал к конденсации. Но в том, что касается ремоделирования хроматина, эффект фосфорилирования как раз противоположный. Возможно, конечно, что эффекты фосфорилирования разных гистонов или даже разных аминокислотных остатков одного гистона могут быть различными вплоть до полной противоположности.

30.13 У белков, модифицирующих хроматин, есть общие аминокислотные мотивы

Ключевые понятия:

- В некоторых белках хроматина есть хромодомен; эффект воздействия таких белков на структуру хроматина может быть активирующим либо репрессирующим.
- Частью активного центра метилтрансфераз служит SET-домен.
- Во многих взаимодействующих с хроматином белках есть бромодомен; он используется для узнавания ацетилированных сайтов в молекулах гистонов.

Наше понимание механизмов контроля структуры хроматина началось с идентификации мутантов, у которых нарушена мозаичность, обусловленная эффектом положения. У *Drosophila* найдено около 30 генов с такими мутациями. Им систематически присваивались названия *Su(var)* (suppressed variegation; если действие дефектного подукта супрессировало мозаичность) или *E(var)* (enhanced variegation; если действие дефектного продукта усиливало эффект мозаичности). Мутациями *Su(var)* отмечены гены, чьи продукты нужны для формирования гетерохроматина. Это, в частности, ферменты, действующие на хроматин (например, гистондеацетилазы), а также белки, которые локализуются в гетерохроматине. Мутации *E(var)* содержатся в генах, чьи продукты нужны для активации экспрессии в других локусах. Так, мутациями *E(var)* промаркированы компоненты комплекса SWI/SNF. Отсюда, кстати, незамедлительно следует, что ремоделирование и другие структурные модификации хроматина помогают контролировать формирование гетерохроматина. О всеобщности этой функциональной связки говорит то, что многие локусы, где имеют место мутации *Su(var)* или *E(var)*, имеют гомологов (с аналогичными функциями продуктов) в геноме дрожжей. Так, у *Schizosaccharomyces pombe* мутации по генам *clr* (cryptic loci regulator) нарушают сайленсинг.

У многих белков Su(var) и E(var) есть общий мотив из 60 аминокислот, соответствующий так называемому хромодомену. Присутствие этого домена в белках обеих групп указывает на то, что он участвует в белок-белковых взаимодействиях с мишенями в составе хроматина. Хромодомен предназначен для привлечения (рекрутинга) белков, которые его содержат, на работу в хроматине. Эта функция выполняется за счет узнавания хромодоменом метилированных лизинов в «хвостах» гистонов (см. разд. 31.3 «Существование гетерохроматина обусловлено взаимодействиями гистонов с другими белками» и разд. 31.4 «Репрессоры и активаторы: белки-антагонисты Polycomb и Trithorax»).

Белок Su(var)3-9 имеет хромодомен, домен, называемый Su(var)3-9, enhancer-of-zeste, Trithorax, сокращенно SET. Его мотив обнаруживается у нескольких белков Su(var). У млекопитающих белок с гомологичным мотивом локализуется в центромерном гетерохроматине и является гистонметилтрансферазой, действующая на ^{9}Lys гистона H3 (см. разд. 30.9 «Метилирование гистонов и метилирование ДНК взаимосвязаны»). SET-домен, как часть его активного центра, по сути, свидетельствует о метилазной активности.

Бромодомен обнаружен у множества белков, взаимодействующих с хроматином, например у гистонацетилтрансфераз. Анализ кристаллической структуры показал, что в бромодомене есть сайт связывания с ацетилированным лизином. Сам по себе бромодомен узнает только очень короткую пептидную последовательность из четырех аминокислот (в числе которых — ацетилированный лизин), так что специфичность к мишени должна зависеть от взаимодействий, в которых участвуют и другие области белка. Кроме гистонацетилтрансфераз бромодомен обнаруживается во многих других белках, взаимодействующих с хроматином, в числе которых — компоненты аппарата транскрипции. Бромодомен, несомненно, используется для узнавания ацетилированных гистонов, так что белки, в которых он есть, задействованы, как правило, в активации генов.

Несмотря на общую закономерность, что активный хроматин ацетилирован, а неактивный — метилирован по гистонам, из этого правила существует ряд исключений. Самое известное исключение — это ацетилирование H4 по ^{12}Lys, специфичное для гетерохроматина.

На «хвосте» одной молекулы гистона может иметь место множество ковалентных модификаций, причем одна модификация может влиять на возможность другой. Фосфорилирование серина по одному положению может быть необходимым условием для ацетилирования лизина по другому положению. На РИС. 30.20 графически изложена ситуация с «хвостом» H3; этот «хвост» может находиться в одном из двух альтернативных состояний. В неактивном хроматине «хвостам» H3 свойственно иметь метил-^{9}Lys, а в активном — ацетил-^{9}Lys и фосфо-^{10}Ser. Каждое из состояний «хвоста» H3 может доминировать на обширных участках хроматина. Фосфорилирование H3 по ^{10}Ser и метилирование H3 по ^{9}Lys взаимно исключаются; это учтено на рисунке при воспроизведении порядка событий (дефосфорилирование по ^{10}Ser предшествует метилированию по ^{9}Lys). Таким образом, посредством ковалентных модификаций по разным аминокислотным положениям, «хвост» гистона H3 может переходить между «активным» и «неактивным» состояниями.

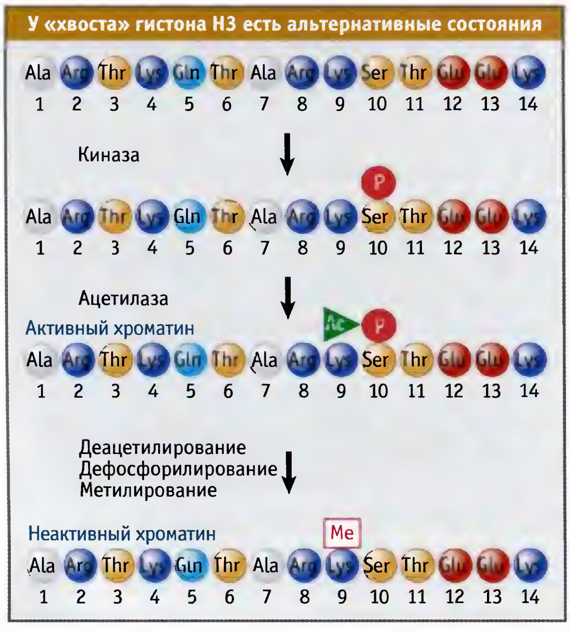


РИС. 30.20 **Различные модификации «хвоста» H3 влияют на активность хроматина**

30.14 Резюме

Гены, чьи регуляторные области имеют стандартную нуклеосомную организацию, обычно не экспрессируются. В отсутствие специальных регуляторных белков, промоторы и другие регуляторные области структурированы гистоновыми октамерами до состояния, в котором они не могут быть активированы. Строго регламентированное позиционирование нуклеосом в окрестностях промоторов можно объяснить потребностью в доступе к важным регуляторным сайтам ДНК. Некоторые факторы транскрипции умеют узнавать ДНК на поверхности нуклеосомы, но все равно, пребывание ДНК в конкретном положении относительно гистонового октамера может быть требованием, необходимым для инициации транскрипции.

Активный хроматин и неактивный хроматин не находятся в равновесии. Чтобы перевести один в другой, нужны события, носящие характер внезапного вмешательства. Белковые комплексы ремоделирования хроматина обладают способностью изменять местополо-

жение гистонов относительно ДНК, используя для этого гидролиз АТФ. Комплексы ремоделирования отличаются большими размерами; их классифицируют по типу АТФазной субъединицы. Самые распространенные из них — это комплексы типа SWI/SNF и типа ISWI. Наиболее часто встречающаяся форма ремоделирования хроматина заключается в смещении одного или нескольких гистоновых октамеров с конкретных последовательностей ДНК: тем самым задается граница для точного либо приоритетного позиционирования прилежащих нуклеосом. Вместо полного исключения нуклеосом, ремоделирование хроматина может свестись к изменениям в положениях нуклеосом, например, за счет скольжения гистоновых октамеров вдоль ДНК.

Ацетилирование гистонов происходит как при репликации, так и при транскрипции; оно может быть необходимо для придания хроматину менее компактной структуры. Некоторые коактиваторы, наводящие взаимосвязь между факторами транскрипции и основной аппарат транскрипции, обладают гистонацетилазной активностью. Репрессоры, напротив, могут быть ассоциированы с гистондеацетилазами. Ферменты, осуществляющие ковалентные модификации гистонов, как правило, специфичны к конкретным аминокислотам конкретных гистонов. Сайты для модификации, используемые наиболее часто, расположены в N-концевых «хвостах» гистонов H3 и H4, которые, проходя между витками ДНК, высовываются из нуклеосом наружу. Активирующие (и репрессирующие) комплексы обычно очень крупные и зачастую обладают несколькими энзиматическими активностями, которые занимаются различными модификациями хроматина. У белков, модифицирующих хроматин, обнаруживается ряд общих мотивов, которым соответствуют домены с определенной функцией. Так, хромодомен обслуживает белок-белковые взаимодействия, бромодомен узнает ацетилированный лизин, а SET-домен является частью активного центра гистонметилтрансферазы.

Модификация «хвостов» гистонов срабатывает как переключатель на запуск реорганизации хроматина. Активация генов, в общем случае, сопровождается ацетилированием «хвостов». Гистонацетилазы можно встретить в составе активирующих комплексов, а гистондеацетилазы — в инактивирующих комплексах. Метилирование гистонов, напротив, сопутствует инактивации генов. Некоторые модификации гистонов могут быть взаимноисключающими или, наоборот, синергичными по отношению к другим модификациям.

Литература

30.2 Хроматин может находиться в альтернативных состояниях

Обзор

Brown, D. D. (1984). The role of stable complexes that repress and activate eukaryotic genes. *Cell* 37, 359–365.

Weintraub, H. (1985). Assembly and propagation of repressed and derepressed chromosomal states. *Cell* 42, 705–711.

Исследование

Bogenhagen, D. F., Wormington, W. M., and Brown, D. D. (1982). Stable transcription complexes of *Xenopus* 5S PHK genes: a means to maintain the differentiated state. *Cell* 28, 413–421.

Workman, J. L. and Roeder, R. G. (1987). Binding of transcription factor TFIID to the major late promoter during *in vitro* nucleosome assembly potentiates subsequent initiation by PHK polymerase II. *Cell* 51, 613–622.

30.3 Ремоделирование хроматина – это активный процесс

Обзор

Becker, P. B. and Horz, W. (2002). ATP-dependent nucleosome remodeling. *Annu. Rev. Biochem.* 71, 247–273.

Felsenfeld, G. (1992). Chromatin as an essential part of the transcriptional mechanism. *Nature* 355, 219–224.

Grunstein, M. (1990). Histone function in transcription. *Annu. Rev. Cell Biol.* 6, 643–678.

Narlikar, G. J., Fan, H. Y., and Kingston, R. E. (2002). Cooperation between complexes that regulate chromatin structure and transcription. *Cell* 108, 475–487.

Peterson, C. L. and Côté, J. (2004). Cellular machineries for chromosomal DNA repair. *Genes Dev.* 18, 602–616.

Tsukiyama, T. (2002). The *in vivo* functions of ATP-dependent chromatin-remodelling factors. *Nat. Rev. Mol. Cell Biol.* 3, 422–429.

Vignali, M., Hassan, A. H., Neely, K. E., and Workman, J. L. (2000). ATP-dependent chromatin-remodeling complexes. *Mol. Cell Biol.* 20, 1899–1910.

Исследование

Cairns, B. R., Kim, Y.-J., Sayre, M. H., Laurent, B. C., and Kornberg, R. (1994). A multisubunit complex containing the SWI/ADR6, SWI2/1, SWI3, SNF5, and SNF6 gene products isolated from yeast. *Proc. Natl. Acad. Sci. USA* 91, 1950–1954.

Cote, J., Quinn, J., Workman, J. L., and Peterson, C. L. (1994). Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. *Science* 265, 53–60.

Gavin, I., Horn, P. J., and Peterson, C. L. (2001). SWI/SNF chromatin remodeling requires changes in DNA topology. *Mol. Cell* 7, 97–104.

Hamiche, A., Kang, J. G., Dennis, C., Xiao, H., and Wu, C. (2001). Histone tails modulate nucleosome mobility and regulate ATP-dependent nucleosome sliding by NURF. *Proc. Natl. Acad. Sci. USA* 98, 14316–14321.

Kingston, R. E. and Narlikar, G. J. (1999). ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. *Genes Dev.* 13, 2339–2352.

Kwon, H., Imbaizano, A. N., Khavari, P. A., Kingston, R. E., and Green, M. R. (1994). Nucleosome disruption and enhancement of activator binding of human SWI/SNF complex. *Nature* 370, 477–481.

Logie, C. and Peterson, C. L. (1997). Catalytic activity of the yeast SWI/SNF complex on reconstituted nucleosome arrays. *EMBO J.* 16, 6772–6782.

Lorch, Y., Cairns, B. R., Zhang, M., and Kornberg, R. D. (1998). Activated RSC-nucleosome complex and persistently altered form of the nucleosome. *Cell* 94, 29–34.

Lorch, Y., Zhang, M., and Kornberg, R. D. (1999). Histone octamer transfer by a chromatin-remodeling complex. *Cell* 96, 389–392.

Peterson, C. L. and Herskowitz, I. (1992). Characterization of the yeast SWI1, SWI2, and SWI3 genes, which encode a global activator of transcription. *Cell* 68, 573–583.

Robert, F., Young, R. A., and Struhl, K. (2002). Genome-wide location and regulated recruitment of the RSC nucleosome remodeling complex. *Genes Dev.* 16, 806–819.

Schnitzler, G., Sif, S., and Kingston, R. E. (1998). Human SWI/SNF interconverts a nucleosome between its base state and a stable remodeled state. *Cell* 94, 17–27.

Tamkun, J. W., Deuring, R., Scott, M. P., Kissinger, M., Pattatucci, A. M., Kaufman, T. C., and Kennison, J. A. (1992). brahma: a regulator of *Drosophila* homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. *Cell* 68, 561–572.

Tsukiyama, T., Daniel, C., Tamkun, J., and Wu, C. (1995). ISWI, a member of the SWI2/SNF2 ATPase family, encodes the 140 кДа subunit of the nucleosome remodeling factor. *Cell* 83, 1021–1026.

Tsukiyama, T., Palmer, J., Landel, C. C., Shiloach, J., and Wu, C. (1999). Characterization of the imitation switch subfamily of ATP-dependent chromatin-remodeling factors in *S. cerevisiae*. *Genes Dev.* 13, 686–697.

Whitehouse, I., Flaus, A., Cairns, B. R., White, M. F., Workman, J. L., and Owen-Hughes, T. (1999). Nucleosome mobilization catalysed by the yeast SWI/SNF complex. *Nature* 400, 784–787.

30.4 Нуклеосомная организация промотора может быть нестандартной

Обзор

Lohr, D. (1997). Nucleosome transactions on the promoters of the yeast GAL and PHO genes. *J. Biol. Chem.* 272, 26795–26798.

Исследование

Cosma, M. P., Tanaka, T., and Nasmyth, K. (1999). Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle and developmentally regulated promoter. *Cell* 97, 299–311.

Kadam, S., McAlpine, G. S., Phelan, M. L., Kingston, R. E., Jones, K. A., and Emerson, B. M. (2000). Functional selectivity of recombinant mammalian SWI/SNF subunits. *Genes Dev.* 14, 2441–2451.

McPherson, C. E., Shim, E.-Y., Friedman, D. S., and Zaret, K. S. (1993). An active tissue-specific enhancer and bound transcription factors existing in a precisely positioned nucleosomal array. *Cell* 75, 387–398.

Schmid, V. M., Fascher, K.-D., and Horz, W. (1992). Nucleosome disruption at the yeast PHO5 promoter upon PHO5 induction occurs in the absence of DNA replication. *Cell* 71, 853–864.

Truss, M., Barstch, J., Schelbert, A., Hache, R. J. G., and Beato, M. (1994). Hormone induces binding of receptors and transcription factors to a rearranged nucleosome on the MMTV promoter *in vitro*. *EMBO J.* 14, 1737–1751.

Tsukiyama, T., Becker, P. B., and Wu, C. (1994). ATP-dependent nucleosome disruption at a heat shock promoter mediated by binding of GAGA transcription factor. *Nature* 367, 525–532.

Yudkovsky, N., Logie, C., Hahn, S., and Peterson, C. L. (1999). Recruitment of the SWI/SNF chromatin remodeling complex by transcriptional activators. *Genes Dev.* 13, 2369–2374.

30.5 Модификация гистонов как ключевой фактор изменения структуры хроматина

Обзор

Jenuwein, T. and Allis, C. D. (2001). Translating the histone code. *Science* 293, 1074–1080.

Исследование

Osada, S., Sutton, A., Muster, N., Brown, C. E., Yates, J. R., Sternglanz, R., and Workman, J. L. (2001). The yeast SAS (something about silencing) protein complex contains a MYST-type putative acetyltransferase and functions with chromatin assembly factor ASF1. *Genes Dev.* 15, 3155–3168.

30.6 Ацетилирование гистонов происходит в двух случаях

Обзор

Hirose, Y. and Manley, J. L. (2000). PHK polymerase II and the integration of nuclear events. *Genes Dev.* 14, 1415–1429.

Verreault, A. (2000). De novo nucleosome assembly: new pieces in an old puzzle. *Genes Dev.* 14, 1430–1438.

Исследование

Akhtar, A. and Becker, P. B. (2000). Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in *Drosophila*. *Mol. Cell* 5, 367–375.

Alwine, J. C., Kemp, D. J., and Stark, G. R. (1977). Method for detection of specific PHKs in agarose gels by transfer to diazobenzyl-oxymethyl-paper and hybridization with DNA probes. *Proc. Natl. Acad. Sci. USA* 74, 5350–5354.

Jackson, V., Shires, A., Tanphaichitr, N., and Chalkley, R. (1976). Modifications to histones immediately after synthesis. *J. Mol. Biol.* 104, 471–483.

Ling, X., Harkness, T. A., Schultz, M. C., Fisher-Adams, G., and Grunstein, M. (1996). Yeast histone H3 and H4 amino termini are important for nucleosome assembly *in vivo* and *in vitro*: redundant and position-independent functions in assembly but not in gene regulation. *Genes Dev.* 10, 686–699.

Shibahara, K., Verreault, A., and Stillman, B. (2000). The N-terminal domains of histones H3 and H4 are not necessary for chromatin assembly factor-1-mediated nucleosome assembly onto replicated DNA *in vitro*. *Proc. Natl. Acad. Sci. USA* 97, 7766–7771.

Turner, B. M., Birley, A. J., and Lavender, J. (1992). Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in *Drosophila* polytene nuclei. *Cell* 69, 375–384.

30.7 Гистонацетилазы ассоциированы с активаторами транскрипции

Исследование

Brownell, J. E. et al. (1996). *Tetrahymena* histone acetyltransferase A: a homologue to yeast Gcn5p linking histone acetylation to gene activation. *Cell* 84, 843–851.

Chen, H. et al. (1997). Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CP/p300. *Cell* 90, 569–580.

Grant, P. A. et al. (1998). A subset of TAFIIs are integral components of the SAGA complex required for nucleosome acetylation and transcriptional stimulation. *Cell* 94, 45–53.

Kingston, R. E. and Narlikar, G. J. (1999). ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. *Genes Dev.* 13, 2339–2352.

Lee, T. I., Causton, H. C., Holstege, F. C., Shen, W. C., Hannett, N., Jennings, E. G., Winston, F., Green, M. R., and Young, R. A. (2000). Redundant roles for the $TF_{II}D$ and SAGA complexes in global transcription. Nature 405, 701–704.

30.8 Гистондеацетилазы ассоциированы с репрессорами транскрипции

Обзор

Richards, E. J., Elgin, S. C., and Richards, S. C. (2002). Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell* 108, 489–500.

Исследование

Ayer, D. E., Lawrence, Q. A., and Eisenman, R. N. (1995). Mad-Max transcriptional repression is mediated by tePHKry complex formation with mammalian homologs of yeast repressor Sin3. *Cell* 80, 767–776.

Kadosh, D. and Struhl, K. (1997). Repression by Ume6 involves recruitment of a complex containing Sin3 corepressor and Rpd3 histone deacetylase to target promoters. *Cell* 89, 365–371.

Schreiber-Agus, N., Chin, L., Chen, K., Torres, R., Rao, G., Guida, P., Skoultchi, A. I., and DePinho, R. A. (1995). An amino-terminal domain of Mxi1 mediates anti-Myc oncogenic activity and interacts with a homolog of the yeast transcriptional repressor SIN3. *Cell* 80, 777–786.

30.9 Метилирование гистонов и метилирование ДНК взаимосвязаны

Обзор

Bannister, A. J., and Kouzarides, T. (2005). Reversing histone methylation. *Nature* 436, 1103–1106.

Richards, E. J., Elgin, S. C., and Richards, S. C. (2002). Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell* 108, 489–500.

Zhang, Y. and Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev.* 15, 2343–2360.

Исследование

Cuthbert, G. L., Daujat, S., Snowden, A. W., Erdjument-Bromage, H., Hagiwara, T., Yamada, M., Schneider, R., Gregory, P. D., Tempst, P., Bannister, A. J., and Kouzarides, T. (2004). Histone deimination antagonizes arginine methylation. *Cell* 118, 545–553.

Fuks, F., Hurd, P. J., Wolf, D., Nan, X., Bird, A. P., and Kouzarides, T. (2003). The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *J. Biol. Chem.* 278, 4035–4040.

Gendrel, A. V., Lippman, Z., Yordan, C., Colot, V., and Martienssen, R. A. (2002). Dependence of heterochromatic histone H3 methylation patterns on the *Arabidopsis* gene DDM1. *Science* 297, 1871–1873.

Johnson, L., Cao, X., and Jacobsen, S. (2002). Interplay between two epigenetic marks. DNA methylation and histone H3 lysine 9 methylation. *Curr. Biol.* 12, 1360–1367.

Lawrence, R. J., Earley, K., Pontes, O., Silva, M., Chen, Z. J., Neves, N., Viegas, W., and Pikaard, C. S. (2004). A concerted DNA methylation/histone methylation switch regulates rPHK gene dosage control and nucleolar dominance. *Mol. Cell* 13, 599–609.

Ng, H. H., Feng, Q., Wang, H., Erdjument-Bromage, H., Tempst, P., Zhang, Y., and Struhl, K. (2002). Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association. *Genes Dev.* 16, 1518–1527.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z. W., Sun, M., Opravil, S., Mechtler, K., Ponting, C. P., Allis, C. D., and Jenuwein, T. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406, 593–599.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., and Casero, R. A. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 119, 941–953.

Tamaru, H. and Selker, E. U. (2001). A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa. Nature* 414, 277–283.

Tamaru, H., Zhang, X., McMillen, D., Singh, P. B., Nakayama, J., Grewal, S. I., Allis, C. D., Cheng, X., and Selker, E. U. (2003). Trimeth-ylated lysine 9 of histone H3 is a mark for DNA methylation in *Neurospora crassa. Nat. Genet.* 34, 75–79.

Wang, Y., Wysocka, J., Sayegh, J., Lee, Y. H., Perlin, J. R., Leonelli, L., Sonbuchner, L. S., McDonald, C. H., Cook, R. G., Dou, Y., et al. (2004). Human PAD4 regulates histone arginine methylation levels via demethylimination. *Science* 306, 279–283.

30.11 Активация промотора представляет собой упорядоченную серию событий

Обзор

Orphanides, G. and Reinberg, D. (2000). PHK polymerase II elongation through chromatin. *Nature* 407, 471–475.

Исследование

Bortvin, A. and Winston, F. (1996). Evidence that Spt6p controls chromatin structure by a direct interaction with histones. *Science* 272, 1473–1476.

Cosma, M. P., Tanaka, T., and Nasmyth, K. (1999). Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle and developmentally regulated promoter. *Cell* 97, 299–311.

Hassan, A. H., Neely, K. E., and Workman, J. L. (2001). Histone acetyltransferase complexes stabilize swi/snf binding to promoter nucleosomes. *Cell* 104, 817–827.

Orphanides, G., LeRoy, G., Chang, C. H., Luse, D. S., and Reinberg, D. (1998). FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell* 92, 105–116.

Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K., Hartzog, G. A., Winston, F., Buratowski, S., and Handa, H. (1998). DSIF, a novel transcription elongation factor that regulates PHK polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes Dev.* 12, 343–356.

30.12 Фосфорилирование гистонов влияет на структуру хроматина

Исследование

Wang, Y., Zhang, W., Jin, Y., Johansen, J., and Johansen, K. M. (2001). The JIL-1 tandem kinase mediates histone H3 phosphorylation and is required for maintenance of chromatin structure in *Drosophila*. *Cell* 105, 433–443.

30.13 У белков, модифицирующих хроматин, есть общие аминокислотные мотивы

Исследование

Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K., and Zhou, M. M. (1999). Structure and ligand of a histone acetyltransferase bromo domain. *Nature* 399, 491–496.

Koonin, E. V., Zhou, S., and Lucchesi, J. C. (1995). The chromo superfamily: new members, duplication of the chromo domain and possible role in delivering transcription regulators to chromatin. *Nucleic Acids Res.* 23, 4229–4233.

Litt, M. D., Simpson, M., Gaszner, M., Allis, C. D., and Felsenfeld, G. (2001). Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science* 293, 2453–2455.

Owen, D. J., OPHKghi, P., Yang, J. C., Lowe, N., Evans, P. R., Ballario, P., Neuhaus, D., Filetici, P., and Travers, A. A. (2000). The structural basis for the recognition of acetylated histone H4 by the bromo domain of histone acetyltransferase Gcn5p. *EMBO J.* 19, 6141–6149.

Turner, B. M., Birley, A. J., and Lavender, J. (1992). Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in *Drosophila* polytene nuclei. *Cell* 69, 375–384.

Наследование эпигенетических эффектов

КРАТКОЕ СОДЕРЖАНИЕ ГЛАВЫ

31.1 Введение

- Эпигенетические эффекты могут обеспечиваться или модификацией нуклеиновой кислоты после ее синтеза, или существованием самовоспроизводящихся белковых структур.

31.2 Гетерохроматин распространяется от места зарождения

- Гетерохроматин зарождается на особой последовательности ДНК, после чего область с неактивной структурой распространяется вдоль хроматинового волокна.
- Гены, находящиеся внутри гетерохроматина, инактивированы.
- Протяженность конкретного неактивного участка в разных клетках отличается, так что инактивация окрестных генов — событие неопределенное и может сказываться в фенотипе как мозаичность, обусловленная эффектом положения.
- Похожий эффект распространения структурных состояний хроматина имеет место в теломерах, а также в локусах, где хранятся молчащие кассеты для детерминации типа спаривания у дрожжей.

31.3 Существование гетерохроматина обусловлено взаимодействиями гистонов с другими белками

- У млекопитающих ключевую роль в формировании гетерохроматина играет белок HP1; он работает, связываясь с метилированным гистоном H3.
- Формирование гетерохроматина у дрожжей инициируется белком Rap1, который связывает ДНК по специфическим последовательностям-мишеням.
- К мишеням Rap1 относятся теломерные повторы, а также сайленсеры в *HML* и *HMR*.
- Rap1 рекрутирует Sir3 и Sir4, которые взаимодействуют с N-концевыми «хвостами» гистонов H3 и H4, а также друг с другом.

31.4 Репрессоры и активаторы: белки-антагонисты Polycomb и Trithorax

- При митозе, белки группы Polycomb (Pc-G) удерживают хроматин в репрессированном состоянии.
- Последовательность ДНК, необходимую для функции Pc-G, называют PRE.
- Последовательность PRE служит организатором, от которого, при участии белков Pc-G, в хроматине идет распространение неактивной структуры.
- Отдельно взятого белка Pc-G, способного связать PRE, пока не найдено.
- Действия белков группы Trithorax противоположны действиям белков Pc-G.

31.5 X-хромосомы претерпевают глобальные изменения

- У самок плацентарных млекопитающих в эмбриогенезе в каждой клетке инактивируется одна, выбранная случайным образом, X-хромосома.
- В исключительных случаях, когда в клетке содержится больше двух X-хромосом, инактивируются все, кроме одной.
- Центр инактивации — это *цис*-действующая область X-хромосомы, оказывающая свое действие в *цис*-манере и являющаяся необходимой и достаточной гарантией того, что из X-хромосом, присутствующих в клетке, активной останется лишь одна.
- Центр инактивации содержит ген *Xist*; он кодирует РНК, которую можно обнаружить только на неактивных X-хромосомах.
- Механизм, препятствующий накоплению РНК *Xist* на активных хромосомах, неизвестен.

31.6 Конденсация хромосомы — функция конденсинов

- Белки SMC являются АТФазами; к ним относят конденсины и когезины.
- Гетеродимер белков SMC ассоциирует с другими субъединицами.
- Конденсины сверхспирализует хроматин за счет того, что вносят в ДНК положительные сверхвитки.
- Конденсины ответственны за конденсацию хромосом при митозе.
- У *C. elegans* особая разновидность конденсинов ответственна за конденсацию X-хромосом, понижающую их активность.

31.7 Поддерживающая метилаза обеспечивает воспроизводство статуса метилирования ДНК

- Большинство метильных групп ДНК обнаружено у цитозинов обеих цепей дублетов CpG.
- Репликация переводит полностью метилированный сайт ДНК в два наполовину метилированных сайта.
- Полуметилированные сайты превращаются в полностью метилированные сайты под действием фермента — поддерживающей метилазы.

31.8 Метилирование ДНК обусловливает импринтинг

- При оплодотворении два аллеля одного и того же гена, отцовский и материнский, могут обладать разными паттернами метилирования.
- Метилирование, как правило, ассоциировано с инактивацией гена.
- Если гены метилированы по-разному (импринтированы), эмбрион выживет при условии, что функциональный (не несущий инактивирующей мутации) аллель достался ему от родителя, у которого он не метилирован.
- Выживаемость гетерозигот по импринтированным генам различается в зависимости от направления скрещивания.
- Импринтируемые гены встречаются в составе кластеров, где их статус может находиться под контролем локальной регуляторной области, а метилирование *de novo* происходит в случае отсутствия предотвращающих метилирование механизмов.

31.9 Импринтинг генов с противоположной половой специфичностью может контролироваться из одного центра на хромосоме

- Метилирование может быть причиной инактивации одних генов и активации других генов.
- Импринтинг происходит в результате метилирования *цис*-действующих сайтов.

31.10 Эпигенетические эффекты могут наследоваться

- Наследование эпигенетических эффектов может обеспечиваться воспроизводимыми постсинтетическими модификациями нуклеиновых кислот либо за счет самоподдерживающихся белковых структур.

31.11 Необычное наследование прионов у дрожжей

- Белок Sup35 в растворимой форме (дикого типа) является фактором терминации трансляции.
- Тот же белок может существовать в альтернативной форме олигомерного агрегата; эта форма Sup35 не пригодна для участия в синтезе белка.
- Присутствие олигомерной формы Sup35 заставляет новосинтезированный Sup35 принимать конформацию, исключающую его активность.
- Переход Sup35 из одной формы в другую происходит под влиянием шаперонов.
- Форма Sup35 дикого типа имеет рецессивный генетический статус *psi⁻*, а мутантная форма — доминантный генетический статус *PSI⁺*.

31.12 Прионы как причина заболеваний млекопитающих

- Белок, которым обусловлено заболевание скрэпи у овец, существует в двух формах; неинфекционная форма дикого типа, PrP^C, расщепляется протеазами, а болезнетворная форма, PrP^{Sc}, устойчива к протеазам.
- Неврологическое заболевание может быть передано от овец мышам при помощи инъекции очищенного белка, приона PrP^{Sc}.
- Чтобы заболеть, мышь-реципиент должна иметь функциональную копию гена *PrP*.
- Прион PrP^{Sc} может поддерживать себя, заставляя новосинтезированный белок PrP принимать форму PrP^{Sc} вместо PrP^C.
- Прион PrP^{Sc} существует в разных вариантах, отличающихся между собой конформацией субъединиц.

31.13 Резюме

31.1 Введение

Ключевые понятия:

- Эпигенетические эффекты могут обеспечиваться или модификацией нуклеиновой кислоты после ее синтеза, или существованием cfvjвоспроизводящихся белковых структур.

Под **эпигенетическим** наследованием подразумевается способность различных состояний генетического материала (имеющих разные фенотипические отображения) быть переданными по наследству без каких-либо изменений в последовательности ДНК. Это означает, что две особи с одной и той же последовательностью ДНК в локусе, от которого исходит этот эффект, могут иметь разные фенотипы. Основная причина феномена заключается в наличии в данном локусе самоподдерживающейся структуры, функционально не зависящей от последовательности ДНК. Способностью закреплять эпигенетические эффекты самим своим существованием обладают следующие типы структур:

- Ковалентная модификация ДНК (метилированное основание).
- Белковая структура, собранная на ДНК.
- Белковый агрегат, контролирующий конформацию новых субъединиц по мере их синтеза.

В каждом из перечисленных случаев эпигенетическое состояние является следствием особенностей функционирования гена (чаще всего, его инактивации), обусловленных данной структурой.

В случае с метилированием ДНК, последовательность, модифицированная таким образом, может не поддаваться транскрипции, в то время как неметилированная последовательность будет экспрессирована. Как наследуется такое состояние, показано на . Один из аллелей имеет последовательность, метилированную по обеим цепям ДНК, в то время как идентичная последовательность другого аллеля не метилирована. Репликация метилированного аллеля порождает наполовину метилированные дочерние дуплексы, которые доводятся до полностью метилированного состояния конститутивно-активным ферментом метилазой. Репликация другого аллеля тоже сохраняет статус его метилирования (то есть отсутствие метильных групп). Если статус метилирования влияет на возможность транскрипции, то два аллеля отличаются по статусу экспрессии, несмотря на их идентичные нуклеотидные последовательности.

Самоподдерживающиеся белковые структуры, проходящие сборку на ДНК, обычно производят репрессирующий эффект тем, что запускают формирование гетерохроматина, внутри которого гены не экспресси-

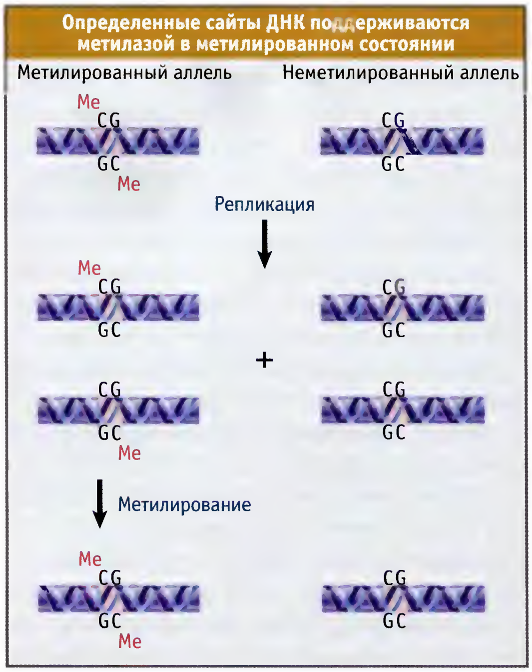


РИС. 31.1 Репликация метилированного сайта ДНК дает сайт, метилированный наполовину (только по только матричной цепи). Специальный фермент, поддерживающая метилаза, узнает в ДНК наполовину метилированные сайты и добавляет метильную группу к нуклеотиду дочерней цепи. Это событие восстанавливает исходную ситуацию: теперь сайт метилирован по обеим цепям. А любой неметилированный сайт после репликации так и остается неметилированным

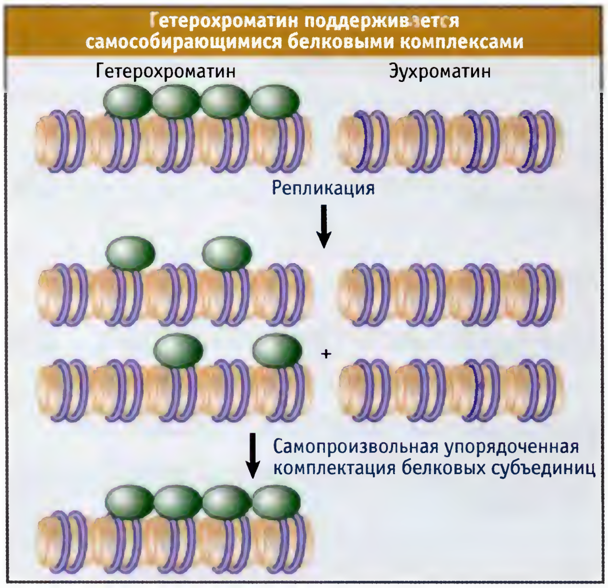


РИС. 31.2 Гетерохроматин образуется благодаря белкам, связанным с гистонами. После репликации ДНК в S-фазе такие белки остаются ассоциированными с каждым из дочерних дуплексов и привлекают новые субъединицы к участию в построении репрессирующего комплекса. Поэтому репрессированное состояние хроматина может передаваться «сквозь» деление клетки

руются. Сохранение областей гетерохроматина зависит от способности определенных белков такой области сохранять свое в ней присутствие и после репликации, а затем вовлекать в свой комплекс в качестве подкрепления еще больше белковых субъединиц. Если при репликации отдельные субъединицы распределятся по дочерним дуплексам случайным образом, то каждый дочерний дуплекс будет по-прежнему помечен белком, хотя плотность его расположения на ДНК будет в два раза ниже, чем до репликации. На РИС. 31.2 показано, что существование эпигенетических эффектов заставляет предполагать у белков, обусловливающих такую ситуацию, способность к самосборке, благодаря которой из материала субъединиц может восстановиться интактный комплекс.

За эпигенетический эффект может отвечать не присутствие белка как такового, но его модификация. Так, в конститутивном гетерохроматине «хвосты» гистонов H3 и H4 обычно не ацетилированы. Если взять центромерный гетерохроматин и подвергнуть ацетилированию, то молчащие в нем гены могут стать активными. Эффект может быть «пронесен» сквозь митоз или мейоз, что характеризует его как эпигенетический, возникший в результате изменения статуса ацетилирования гистонов.

Независимые белковые агрегаты — **прионы** — обусловливают эпигенетические эффекты тем, что, связывая белок, препятствуют выполнению его нормальной функции. Такой белковый агрегат, с момента, как он сформировался, заставляет новосинтезированные белковые субъединицы примыкать к себе, приняв неактивную конформацию.

31.2 Гетерохроматин распространяется от места зарождения

Ключевые понятия:

- Гетерохроматин зарождается на особой последовательности ДНК, после чего область с неактивной структурой распространяется вдоль хроматинового волокна.
- Гены, находящиеся внутри гетерохроматина, инактивированы.
- Протяженность конкретного неактивного участка в разных клетках отличается, так что инактивация окрестных генов — событие неопределенное и может сказываться в фенотипе как мозаичность, обусловленная эффектом положения.
- Похожий эффект распространения структурных состояний хроматина имеет место в теломерах, а также в локусах, где хранятся молчащие кассеты для детерминации типа спаривания у дрожжей.

В интерфазных ядрах кроме эухроматина есть и гетерохроматин, который по степени своей конденсированности близок к митотическим хромосомам. Гетерохроматин инертен. В продолжение интерфазы он остается конденсированным, и транскрипция в нем полностью подавлена. Гетерохроматн реплицируется в поздней S-фазе и может локализоваться на периферии ядра. Центромерный гетерохроматин организуется, в типичном случае, на после-



РИС. 31.3 У мух *Drosophila* цвет глаз зависит от функции гена *white*. Можно наблюдать интересный фенотипический эффект: мозаичность, обусловленную положением гена (эффект положения). Клетки, в которых *white* инактивировался, подпав под экспансию гетерохроматина, образуют в глазу мухи вкрапления белого цвета — на красном фоне активного *white*. Насколько ярко выражен фенотип, зависит от расстояния между гетерохроматином и сайтом интеграции гена *white*. Фотография любезно предоставлена Steven Henikoff, Fred Hutchinson Cancer Research Center

довательностях сателлитной ДНК; в то же время сказать, что формирование гетерохроматина строго продиктовано последовательностью ДНК, нельзя. Если какой-нибудь активный ген расположить по-соседству с гетерохроматином (либо за счет хромосомной транслокации, либо путем трансфекции и интеграции), этот ген в своем новом положении может стать неактивным, что свидетельствует о его переходе в состояние гетерохроматина.

Подобная инактивация носит эпигенетический характер (см. разд. 31.10 «Эпигенетические эффекты могут наследоваться»). Она может составлять предмет различий между индивидуальными клетками в организме животного, что приводит к феномену мозаичности, обусловленной **эффектом положения** (position effect variegation, PEV); суть эффекта состоит в том, что клетки, идентичные генетически, обладают, тем не менее, разными фенотипами. У *Drosophila* этот феномен описан довольно подробно. На РИС. 31.3 показана муха, в глазу которой наблюдается как раз такая мозаичность. Некоторые участки глаза бесцветные, другие — красные. Различие обусловлено тем, что в одних клетках ген *white* из-за близости к гетерохроматину оказался инактивированным, в других же — остался активным.

Разъяснению молекулярной природы этого явления посвящен РИС. 31.4. Эффект инактивации, исходя от гетерохроматина, распространяется по смежным областям. Область такого распространения мозаична: в некоторых клетках оно заходит достаточно далеко для того, чтобы инактивировать близлежащий ген (например, *white*), в других клетках до этого не доходит. Причем случается это на определенном этапе эмбрионального развития, и с этого

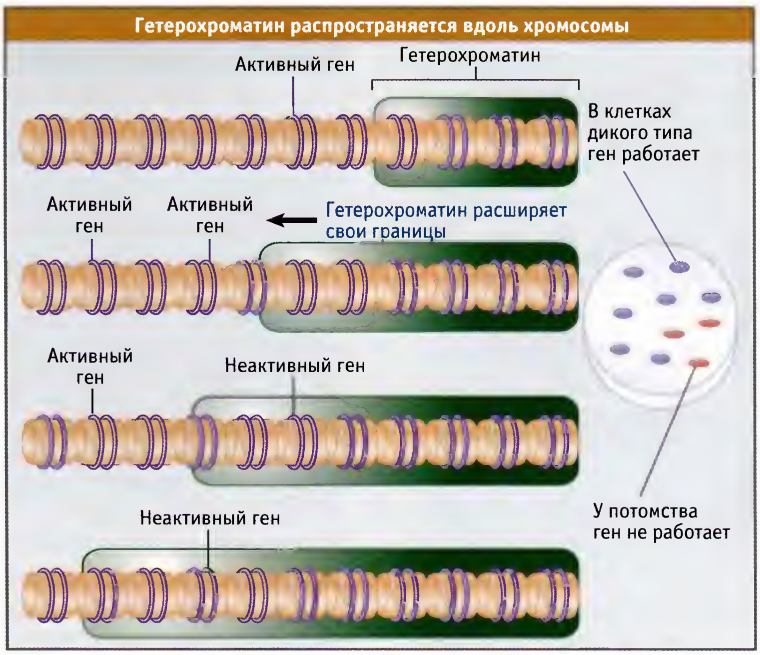


РИС. 31.4 Расширение границ гетерохроматина приводит к инактивации генов. Вероятность инактивации для конкретного гена зависит от расстояния (в пн ДНК) между этим геном и областью гетерохроматина

момента состояние близлежащего гена (активное или неактивное) является уже свойством клеточной линии, поскольку наследуется дочерними клетками при каждом последующем делении. Клетки, ведущие свое происхождение от клетки-предка с инактивированным геном, образуют вкрапления с фенотипом потери функции в массиве ткани с нормальным фенотипом. Так, инактивация гена *white* выражается в отсутствии пигмента, в результате чего в глазу мухи появляются бесцветные вкрапления.

Чем ближе к гетерохроматину расположен ген, тем больше вероятность его инактивации. Формирование гетерохроматина может идти в два приема: сначала, на специально предназначенной для этого последовательности ДНК, происходит нуклеация гетерохроматина, а затем неактивная структура хроматина распространяется вдоль хроматинового волокна. Область распространения точно не задана; размах экспансии может определяться стихийно, под влиянием таких параметров, как лимитирующие количества белковых компонентов. Одним из факторов, которые могут влиять на процесс растекания, является статус активности промоторов данной области; активный промотор может приостановить растекание.

Эффект **теломерного замолкания (сайленсинга)** у дрожжей имеет аналогичную природу. Гены, транслоцированные в окрестности теломеры, проявляют случайную потерю активности; эффект обусловлен растеканием неактивного состояния хроматина, распространяющимся от теломер.

У дрожжей существует и другая форма сайленсинга. Тип спаривания дрожжевой клетки, как известно, определяется активностью локуса *MAT*, однако в клеточном геноме содержится еще два экземпляра последовательностей (кассет), способных диктовать тип спаривания. Они находятся в локусах *HML* и *HMR* и удерживаются клеткой в неактивной форме. Молчащие локусы *HML* и *HMR* имеют ряд свойств, общих с гетерохроматином, так что их можно расценивать как миниатюрные области гетерохроматина (см. разд. 19.22 «Молчащие кассеты в *HML* и *HMR* репрессированы»).

Инактивация хроматина происходит за счет добавления белков к волокну, состоящему из нуклеосом. Инактивация может быть следствием множества разных событий, в том числе, конденсации хроматина (которая делает его недоступным для аппарата экспрессии генов), а также появления белков, напрямую блокирующих доступ к регуляторным сайтам или ингибирующих транскрипцию.

Две системы инактивации хроматина довольно хорошо изучены на молекулярном уровне: одна, с участием белка HP1, у млекопитающих; другая, с участием комплекса SIR, у дрожжей. Хотя между компонентами этих двух систем нет детального сходства, общий принцип реакций у них один и тот же: в качестве точек контакта с хроматином выступают N-концевые «хвосты» нуклеосомных гистонов.

Белок HP1 (heterochromatin protein 1) — это один из самых важных белков Su(var). У *Drosophila* его сначала идентифицировали как белок, локализованный в гетерохроматине: это выяснилось при окрашивании политенных хромосом антителами к нему. Впоследствии оказалось, что этот белок — продукт гена *Su(var)2-5*. Его гомолог у дрожжей *Schizosaccharomyces pombe* кодируется геном *swi6*. Белок, исходно обозначенный как HP1, теперь называют HP1α, потому что с тех пор было идентифицировано еще два родственных белка, HP1β и HP1γ.

У HP1 есть хромодомен, расположенный недалеко от N-конца; другой похожий домен (chromo-shadow domain, что можно перевести с английского как «теневой хромодомен») соответствует С-концевой области полипептида (см. рис. 31.6). Важность хромодомена HP1 доказывается тем, что множество функционально значимых мутаций сконцентрировано именно в нем.

Мутация в деацетилазе, действующей на ацетил-^{14}Lys гистона H3, препятствует метилированию H3 по ^{9}Lys. Гистон H3, метилированный по ^{9}Lys, связывает HP1 по хромодомену. Основываясь на этих фактах, можно построить модель инициации формирования гетерохроматина, изображенную на РИС. 31.5. Сначала приходит деацетилаза и убирает ацетильную группу с ^{14}Lys; затем наступает очередь метилазы SUV39H1: действуя на «хвост» H3, она создает метилированный сигнал, с ко-

31.3 Существование гетерохроматина обусловлено взаимодействиями гистонов с другими белками

Ключевые понятия:

- У млекопитающих ключевую роль в формировании гетерохроматина играет белок HP1; он работает, связываясь с метилированным гистоном H3.
- Формирование гетерохроматина у дрожжей инициируется белком Rap1, который связывает ДНК по специфическим последовательностям-мишеням.
- К мишеням Rap1 относятся теломерные повторы, а также сайленсеры в *HML* и *HMR*.
- Rap1 рекрутирует Sir3 и Sir4, которые взаимодействуют с N-концевыми «хвостами» гистонов H3 и H4, а также друг с другом.

Метилирование гистона H3 позволяет ему связаться с белком HP1

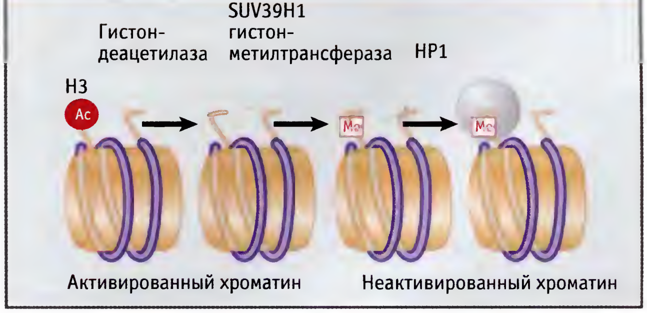


РИС. 31.5 Белок SUV39H1 является гистонметилтрансферазой, действующей на ^{9}Lys гистона H3. С метилированным H3 связывается белок HP1

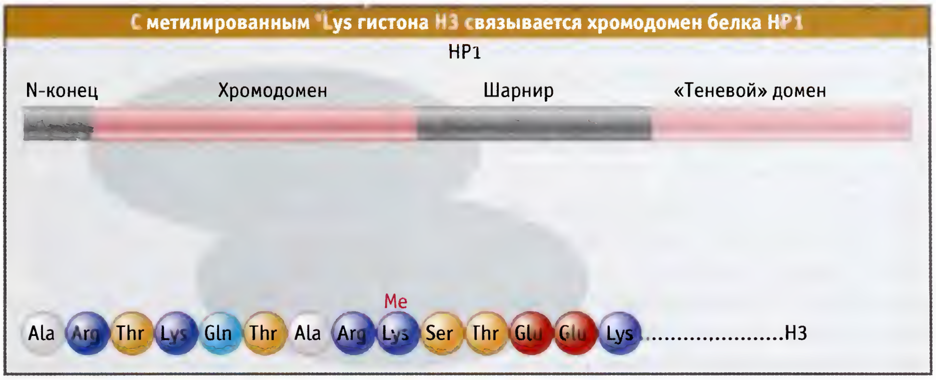


РИС. 31.6 При метилировании у гистона H3 образуется сайт связывания с HP1

торым свяжется HP1. На РИС. 31.6 взаимодействие HP1 с «хвостом» H3 разобрано крупным планом: оно происходит, соответственно, между хромодоменом и метилированным лизином. Это взаимодействие работает как переключатель на формирование гетерохроматина. Как показано на РИС. 31.7, неактивная область хроматина может быть расширена благодаря притоку новых молекул HP1, способных взаимодействовать друг с другом.

У дрожжей любой сайленсинг опирается на один и тот же набор генетических локусов, что говорит о наличии общего механизма. Мутации по любому локусу из этого набора не только приводят к активированному состоянию *HML* и *HMR*, но и облегчают инактивацию генов, интегрированных поблизости от теломерного гетерохроматина. Следовательно, продукты этих локусов работают на поддержание неактивного состояния разных типов гетерохроматина.

На РИС. 31.8 предложена модель функционирования этих белков. Лишь один из них, Rap1, проявляет специфичность к определенной последовательности ДНК. Rap1 связывает ДНК по $C_{1-3}A$-повторам в теломерах, а также по элементам-сайленсерам, действующим в *цис*-манере и необходимым для репрессии *HML* и *HMR*. С Rap1 взаимодействуют белки Sir3 и Sir4, взаимодейству-

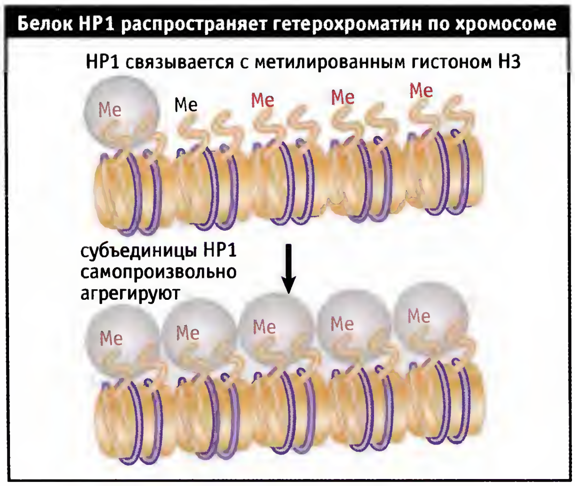


РИС. 31.7 Связывание HP1 с метилированным гистоном H3 стимулирует посадку и агрегацию последующих субъединиц HP1 на череде нуклеосом. Теоретически связывание хроматином одной молекулы HP1 может повлечь за собой инактивацию целой области генома

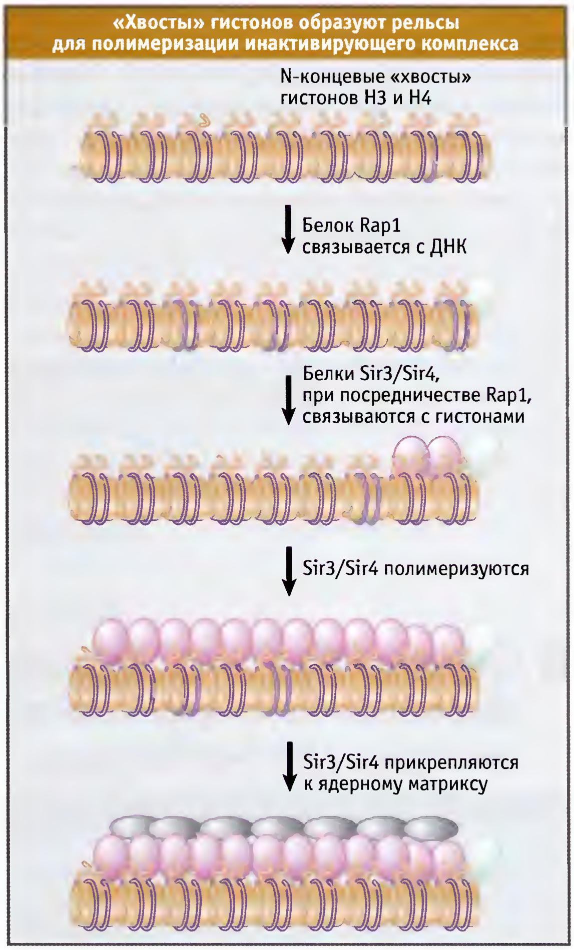


РИС. 31.8 Образование гетерохроматина инициируется ДНК-связывающим белком Rap1. Белки Sir3/Sir4 связывают Rap1 и одновременно – гистоны H3 и H4. Комплекс Sir3/Sir4 полимеризуется вдоль волокон хроматина; возможно, он крепит теломеры хромосом к ядерному матриксу

ющие также между собой (и, возможно, функционирующие как гетеромультимер). Комплекс Sir3/Sir4 взаимодействует с N-концевыми «хвостами» гистонов H3 и H4. На самом деле, первым индикатором того, что гистоны могут принимать непосредственное участие в формировании гетерохроматина, послужило открытие, что мутации, приводящие к отмене сайленсинга по *HML* и *HMR*, приходятся на гены, кодирующие гистоны H3 и H4.

Роль Rap1 в формировании гетерохроматина, в буквальном смысле, ключевая: он узнает последовательности ДНК, с которых стартует весь процесс. Rap1 рекрутирует Sir3/Sir4, и они напрямую взаимодействуют с гистонами H3 и H4. Как только Sir3/Sir4 связались с гистонами H3 и H4, может быть запущена полимеризация этого комплекса (так называемого «комплекса сайленсинга») с эффектом распространения по хроматиновому волокну. Область, охваченная комплексом сайленсинга инактивирetncz: либо из-за ингибирующего действия обкладки из Sir3/Sir4, либо из-за того, что их связывание с гистонами H3/H4 индуцирует какие-то дополнительные структурные изменения. Чем лимитировано растекание комплекса, не известно. Своей C-концевой областью Sir3 напоминает ламины (белки ядерного матрикса) и может отвечать за привязывание гетерохроматина к периферии ядра.

Похожая серия событий оформляет молчащие участки в *HML* и *HMR* (см. также разд. 19.22 «Молчащие кассеты в *HML* и *HMR* репрессированы»). В запуске формирования комплекса участвуют три белка, специфично связывающих определенные последовательности ДНК. Это Rap1, фактор транскрипции Abf1 и ORC (origin replication complex; «комплекс репликации ориджина»). С одним из этих белков, совершивших свою посадку на ДНК, связывается Sir1, который рекрутирует Sir2, Sir3 и Sir4 для формирования репрессивной структуры (комплекса сайленсинга). Sir2 работает как гистондеацетилаза; реакция деацетилирования необходима для поддержания связи между комплексом SIR и хроматином.

Как (по какому механизму) комплексы сайленсинга подавляют активность хроматина? Например, такой комплекс мог бы конденсировать хроматин, чтобы регуляторные белки не находили в нем своих мишеней. Для простоты предположим, что присутствие комплекса сайленсинга несовместимо с присутствием факторов транскрипции и РНК-полимеразы. Природа несовместимости может быть в том, что комплекс сайленсинга препятствует связыванию факторов транскрипции, прямо (закрывая собой ДНК) или косвенно (например, блокируя ремоделирование хроматина). Однако ситуация может быть и не такой простой, поскольку и факторы транскрипции, и РНК-полимераза на промоторах в молчащем хроматине все же присутствуют. Их присутствие может означать, что комплекс сайленсинга препятствует, скорее, работе факторов, чем их связыванию как таковому. На самом деле, между активаторами генов и репрессирующими эффектами в хроматине вполне может существовать конкуренция, в рамках которой активация промотора ингибирует растекание комплекса сайленсинга по хроматиновому волокну.

Формирование гетерохроматина у дрожжей *S. pombe* обусловлено комплексом сайленсинга, в котором, помимо белков, содержится несколько молекул РНКi (см. разд. 13.10, «РНК-интерференция — причина молчания генов»). Эти молекулы РНКi являются продуктами транскрипции центромерных повторов, нарезанными на более мелкие фрагменты. Комплекс также содержит белки, гомологичные тем, что участвуют в формировании гетерохроматина у других организмов: например, белок Argonaute, который участвует в доставке ремоделирующего комплекса RISC (РНК-induced silencing complex; «комплекс для РНК-индуцированного сайленсинга») к его мишеням в хроматине. Компоненты РНКi в составе комплекса отвечают за его локализацию в центромерах. Связавшись там с хроматином, комплекс способствует двойному метилированию (диметилированию) гистона H3 ферментом гистонметилтрансферазой.

В центромерном хроматине формируется еще одна специализированная структура. Ключом к ее природе у *Saccharomyces cerevisiae* служит фенотип мутантов по *cse4* с дефектами в структуре центромер. Белок Cse4 похож на гистон H3. Еще один родственный им белок (с гомологичной аминокислотной последовательностью), белок CENP-A (centromeric protein A; «центромерный белок A»), есть у млекопитающих. Генетические взаимодействия между мутациями в *cse4* и *CDE-II*, а также между мутациями в *cse4* и гене гистона H4, заставляют предполагать, что на основе комплекса Cse4-H4 может сформироваться гистоновый октамер, к которому затем прикрепляются центромерные комплексы CBF1 и CBF3, чтобы образовать центромеру.

Образовавшись, центромера может принять участие в формировании гетерохроматина своей области. В клетках человека, белок CENP-B, специфично локализованный в центромере, выступает инициатором модификаций гистона H3: деацетилирования по ^{9}Lys и ^{14}Lys, с последующим метилированием по ^{9}Lys. Эти модификации запускают ассоциацию H3 со Swi6, что ведет к формированию гетерохроматина.

31.4 Репрессоры и активаторы: белки-антагонисты Polycomb и Trithorax

Ключевые понятия:

- При митозе белки группы Polycomb (Pc-G) удерживают хроматин в репрессированном состоянии.
- Последовательность ДНК, необходимую для функции Pc-G, называют PRE.
- Последовательность PRE служит организатором, от которого, при участии белков Pc-G, в хроматине идет распространение неактивной структуры.
- Отдельно взятого белка Pc-G, способного связать PRE, пока не найдено.
- Действия белков группы Trithorax противоположны действиям белков Pc-G.

Итак, хроматин может находиться в состоянии специфической репрессии; хорошим тому примером служит гетерохроматин. Другой пример был открыт в связи с поведением гомеозисных генов у *Drosophila*; их изучение привело к идентификации белкового комплекса, способного поддерживать определенные гены в репрессированном состоянии. Оказалось, что мутантам *Pc* свойственны трансформации клеточных типов, также присущие фенотипам мутаций с обретением функции в генах *Antennapedia* (*Antp*) или *Ultrabithorax*, вследствие того, что у мутантов по *Pc* эти гены экспрессируются в тканях, где обычно они репрессированы. Поэтому *Pc* была приписана роль в регуляции тренскрипции. Ген *Pc* и в самом деле оказался прототипом целого класса генов, расположенных в ~15 локусах. Класс был назван *Pc-group* (*Pc-G*); мутации в этих генах имеют, в общем, один и тот же результат: дерепрессию гомеозисных генов, что говорит о наличии у всех генов класса *Pc-G* общей регуляторной функции.

Белки Pc-G функционируют в составе обширных комплексов. Один из таких комплексов, PRC1 (Polycomb-repressive complex) содержит сам Pc, несколько других белков Pc-G, а также пять факторов транскрипции общего назначения. Другой комплекс, Esc-E(z), содержит Esc, E(z), другие белки Pc(G), гистонсвязывающий белок, а также гистондеацетилазу. Белок Pc имеет хромодомен, который связывается с метилированным гистоном H3, а белок E(z) является метилтрансферазой, действующей на H3. Такое сочетание отдельных функций поддерживает прямую взаимосвязь между ремоделированием хроматина и его репрессией. Первоначально представление о такой взаимосвязи возникло на основании исследования мутаций в другом гене *Drosophila*: *brahma* (гомолог дрожжевого *Swi2*), кодирующем компонент ремоделирующего комплекса SWI/SNF. Потеря его функции супрессирует мутации в *Polycomb*.

Плейотропность мутаций по *Pc* (многообразие фенотипических эффектов) свидетельствует о том, что Pc является ядерным белком. Его можно встретить в 80 сайтах политенных хромосом (в числе этих сайтов находится и ген *Antp*). На тех же полосах политенных хромосом, что и Pc, можно встретить еще один белок, продукт другого гена из класса *Pc-G*, *polyhomeotic*. Оба белка можно осадить при помощи антител, специфичных к одному из них, по методике иммунопреципитации. Они пребывают в составе единого белкового комплекса, очень массивного (~2,5 × 10^6 Да), включающего другие белки (всего от 10 до 15 полипептидов). Соответствие между этими белками и продуктами ~30 генов *Pc-G* еще предстоит установить (или убедиться в его отсутствии). Возможно, что продукты некоторых генов *Pc-G* формируют основу репрессивного комплекса, после чего к ним присоединяются некоторые другие белки, придающие всему комплексу специфичность.

В обычном понимании, белки Pc-G не являются репрессорами: первоначальный паттерн экспрессии генов-мишеней задается не ими. В отсутствие белков Pc-G, если некоторые гены сначала репрессированы (в соответствии с позицией, которую клетка занимает в теле зародыша), то позднее по ходу развития организма репрессия сходит на нет. Это говорит о том, что белки Pc-G каким-то образом опознают хроматин в установившемся состоянии репрессии, и затем оказывают действие, направленное на

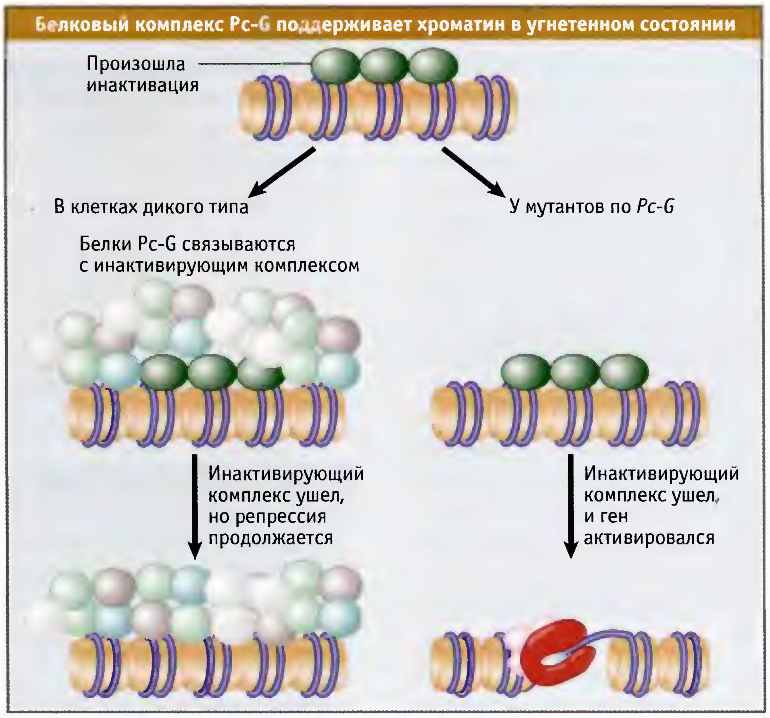


РИС. 31.9 Белки Pc-G не инициируют репрессию генов в хроматине, но ответственны за ее поддержание

поддержание этого состояния (вопреки грядущей череде клеточных циклов, с их репликациями и делениями) во всех клетках-потомках. На рис. 31.9 изображена модель, согласно которой белки Pc-G связывают хроматин совместно с репрессором, но продолжают там оставаться, когда репрессора уже нет. Эта мера необходима, чтобы поддерживать репрессию, и если белков Pc-G на месте не оказалось, то ген, в конце концов, будет активирован.

Участок ДНК, достаточный для того, чтобы ответ на наличие продуктов генов Pc-G состоялся, называют PRE (*Polycomb* response element). Этот участок определяется экспериментально, методом подбора, по способности поддерживать репрессию поблизости от себя, во все продолжение периода развития организма. Стандартный подход заключается во внедрении PRE поблизости от репортерного гена, находящегося под контролем энхансера, который в раннем развитии репрессируется: после этого наблюдают за экспрессией гена-репортера у клеточного потомства; если PRE эффективен, экспрессии не будет.

Элемент PRE, длиной ~10 тпн, укомплектован сложной белковой структурой. Белков, обладающих сродством к сайтам ДНК в пределах PRE, было обнаружено два, но могут быть и другие. Если какой-нибудь локус стал объектом репрессии со стороны Pc-G, репрессирующий комплекс оккупирует область ДНК значительно превосходящую PRE. Так, белок Pc можно встретить на протяжении нескольких тпн с каждой стороны от PRE. Это говорит о том, что PRE может служить организатором нуклеации в хроматине структурного состояния, определяемого белками Pc-G, и, соответственно, центром, из которого идет распространение этого состояния. Такую модель подкрепляют наблюдения за эффектами, подобным мозаичности, обусловленной эффектом положения (см. рис. 31.4): ген, расположенный близко к локусу, репрессия которого поддерживается Pc-G, может стать наследуемо инактивированным в некоторых, но не во всех, клетках. Эксперименты с внесением искусственных сшивок *in vivo* показали, что белок Pc распределен по обширным неактивным областям генного комплекса *Bithorax*, но при этом исключен из участков, содержащих активные гены. Гипотеза о кооперативных взаимодействиях белковых молекул внутри мультимерного комплекса подтверждается существованием в *Pc* мутаций, которые не только меняют распределение белка Pc в ядре, но также лишают продукты других генов класса *Pc-G* способности локализоваться в ядре. То, что роль белков Pc-G состоит именно в поддержании, но не инициации, репрессии, должно означать, что формирование комплекса в PRE зависит также от локального статуса экспрессии генов.

На сегодняшний день, связывание белков Pc-G в PRE описывается рабочей моделью, основанной на свойствах индивидуальных белков. Согласно этой модели, белки Pho и Pho1 связываются в PRE со своими последовательностями-мишенями, после чего комплекс Pho/Pho1 рекрутирует Esc-E(z), который, при помощи своей метилтрансферазной активности, метилирует ^{27}Lys гистона H3. Это создает сайт связывания для PRC1, поскольку хромодомен Pc связывается с метилированным лизином. Комплекс PRC1 переводит хроматин в более компактную структуру; каждая единица такого комплекса затрудняет доступ факторов к примерно трем нуклеосомам.

Хромодомен, на самом деле, был первоначально идентифицирован как участок гомологии между Pc и другим белком, компонентом гетерохроматина HP1. Связывание хромодомена Pc с ^{27}Lys гистона H3 аналогично тому, как HP1 использует свой хромодомен, чтобы связаться с ^{9}Lys. Мозаичность возникает из-за вариабельного распространения инактивации от конститутивного гетерохроматина, поэтому весьма вероятно, что хромодомены HP1 и Pc используются сходным образом, чтобы индуцировать формирование гетерохроматина или другой неактивной структуры (см. разд. 30.13 «У белков, модифицирующих хроматин, есть общие аминокислотные мотивы»). Таким образом, модель подразумевает существование схожих механизмов для репрессии индивидуальных локусов и для образования гетерохроматина.

Белки группы Trithorax (продукты генов *trxG*) создают эффект, противоположный эффекту белков Pc-G: они поддерживают гены в активном состоянии. В механизме действия белков обеих групп может существовать некоторое сходство, поскольку мутации по ряду локусов мешают функционировать как белкам Pc-G, так и белкам trxG; это означает, что белки обеих групп, в своей функции, могут опираться на партнерство одних и тех же компонентов хроматина. Фактор GAGA, кодируемый геном *trithorax-like*, имеет сайт связывания в PRE. Оказалось, что сайты, по которым Pc и GAGA связывают ДНК, совпадают.

Вероятно, фактор GAGA нужен для того, чтобы активирующие факторы, в том числе белки trxG, могли связаться с ДНК. Нужен ли он также белкам Pc-G, чтобы связаться с ДНК и оказать репрессирующее действие, пока не ясно, но такая модель потребовала бы участия, кроме GAGA, дополнительного фактора, определяющего, какой тип комплекса (нацеленного на поддержание активного либо репрессированного состояния хроматина) будет собран на этом месте.

31.5 X-хромосомы претерпевают глобальные изменения

Ключевые понятия:

- У самок плацентарных млекопитающих в эмбриогенезе в каждой клетке инактивируется одна, выбранная случайным образом, X-хромосома.
- В исключительных случаях, когда в клетке содержится больше двух X-хромосом, инактивируются все, кроме одной.
- Центр инактивации — это *цис*-действующая область X-хромосомы, оказывающая свое действие в *цис*-манере и являющаяся необходимой и достаточной гарантией того, что из X-хромосом, присутствующих в клетке, активной останется лишь одна.
- Центр инактивации содержит ген *Xist*; он кодирует РНК, которую можно обнаружить только на неактивных X-хромосомах.
- Механизм, препятствующий накоплению РНК *Xist* на активных хромосомах, неизвестен.

Существование полов (собственно, то обстоятельство, что пол животного задается различиями по половым хромосомам) ставит перед системами регуляции экспрессии генов довольно интересную и весьма насущную проблему. Она заключается в отличии по числу Х-хромосом. Если гены, сцепленные с Х-хромосомой (то есть расположенные на Х-хромосоме), будут экспрессироваться с равными интенсивностями у особей обоих полов, то в клетках тела самки количество продуктов таких генов окажется вдвое больше, чем у самцов. О том, как важно избежать такой ситуации, говорит само наличие явления **компенсации доз** генов и многообразие его механизмов среди разных групп животных. Суть явления состоит в выравнивании интенсивности экспрессии генов, расположенных на Х-хромосоме, между полами. Механизмы такого выравнивания у разных видов животных обобщает таблица на РИС. 31.10.

- У млекопитающих одна из двух Х-хромосом в клетках самки инактивируется полностью. В результате у самок активна только одна Х-хромосома, что эквивалентно ситуации у самцов. Гены активной Х-хромосомы самок и единственной Х-хромосомы самцов экспрессируются с равной интенсивностью.
- Экспрессия генов единственной Х-хромосомы самцов *Drosophila* вдвое превосходит по интенсивности экспрессию генов на каждой из Х-хромосом самок.
- Экспрессия генов на каждой из Х-хромосом самки *Caenorhabditis elegans* вдвое слабее, чем экспрессия тех же генов на одиночной Х-хромосоме самца.

Все перечисленные механизмы компенсации доз генов объединяет то, что *мишенью для регуляции служит хромосома в целом*. То есть имеет место глобальное изменение, затрагивающее все промоторы на хромосоме, причем в количественной манере. Как это происходит, более или менее внятно описано для случая инактивации Х-хромосомы у самок млекопитающих, когда материал целой хромосомы переходит в состояние гетерохроматина.

Двумя неотъемлемыми свойствами гетерохроматина являются его конденсированность и неактивность. Гетерохроматин, как известно, можно подразделить на два типа:

- **Конститутивный гетерохроматин** содержит последовательности ДНК, не несущих кодирующей функции. Их типичный компонент — сателлитные ДНК, которые часто обнаруживаются в центромерах. Такие участки непременно и всегда гетерохроматические в силу своей особой природы.
- **Факультативный гетерохроматин** может быть формой существования целых хромосом, которые в какой-нибудь из клеточных линий не активны, зато в клетках другой линии их гены прекрасно экспрессируются. Типичный пример — Х-хромосома млекопитающих. Неактивная Х-хромосома гетерохроматична, в то время как активная Х-хромосома является частью эухроматина. Таким образом, *идентичные последовательности ДНК оказываются в двух противоположных состояния хроматина*. Как только неактивное состояние установилось, оно начинает наследоваться клеточным потомством. Это пример эпигенетической наследственности, поскольку наследование идет вне зависимости от последовательности ДНК.

Основа современных представлений о ситуации с Х-хромосомой у самок млекопитающих была заложена еще в 1961 г., в форме **гипотезы об одиночной Х**. Самки мышей, гетерозиготные по мутациям, обусловливающим окраску шерсти в манере, сцепленной с Х-хромосомой, имеют мозаичный фенотип: шерсть на их теле окрашена пестро, участки цвета дикого типа перемежаются участками «мутантного» цвета. На РИС. 31.11 разъяснено, что такой фенотипический эффект можно отнести на счет *случайно инактивации одной из двух Х-хромосом в каждой клетке небольшой популяции клеток-предшественников*. Клетки, в которых оказалась инактивирована Х-хромосома, несущая ген дикого типа, посредством делений дают начало потомству (в смысле, популяции клеток), в котором экспрессируется только мутантный аллель (расположенный на активной хромосоме). Клетки, происходящие от клетки-предшественника, в которой была инактивирована другая хромосома, несут в себе активный ген дикого типа. В случае с окраской шерсти, клетки, ведущие свое происхождение от конкретной клетки-предшественника, остаются вместе, формируя таким образом одноцветный лоскут в паттерне окраски животного; это обстоятельство делает мозаичность наглядной. В других случаях клетки экспрессируют один из двух аллелей, расположенных в их Х-хромосомах, в индивидуальном порядке. Так, у гетерозигот по одному из локусов Х-хромосомы, *G6PD*, каждая отдельно взятая красная кровяная клетка, эритроцит, будет экспрессировать какой-нибудь один из двух аллелей.

Инактивация одной Х-хромосомы по выбору случайным образом имеет место у плацентарных млекопитающих. У сумчатых этот выбор направленный: инактивируется всегда Х-хромосома, полученная от отца.

Инактивация Х-хромосомы подчиняется **правилу n–1**: сколько бы ни было в клетке Х-хромосом, инакти-



РИС. 31.10 Животные используют различные стратегии выравнивания экспрессии генов Х-хромосомы в мужском и женском организмах

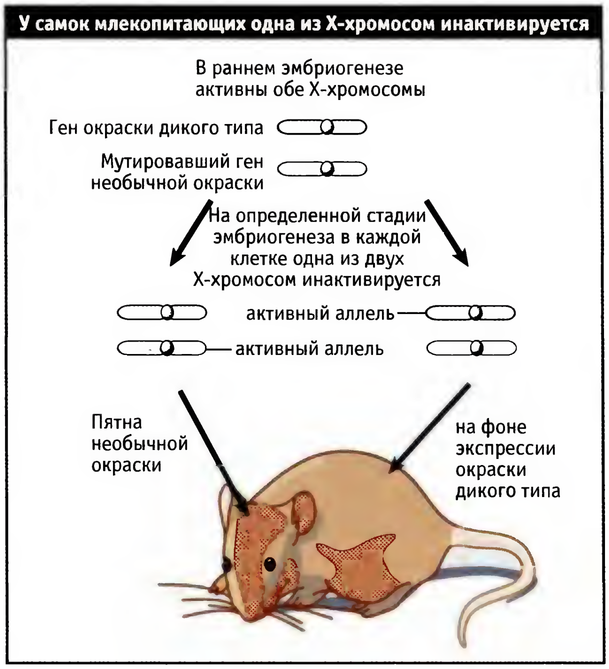


РИС. 31.11 Мозаичность, сцепленная с полом, проявляется у самок млекопитающих. Это эффект *случайного выбора* одной из X-хромосом в каждом субъекте небольшой популяции гетерозиготных клеток-предшественников. Клетки, у которых осталась активной X-хромосома с аллелем дикого типа, дадут дикий фенотип; другие клетки (с активной X-хромосомой, несущей мутацию) будут проявлять мутантный фенотип

вированы будут все, кроме одной. У нормальных самок X-хромосом, конечно же, две, но в тех редких случаях, когда неправильное расхождение хромосом приводит к генотипу с тремя или более X-хромосомами, активной останется лишь одна из них. Это обстоятельство наводит на общую модель, в соответствии с которой с одной и только одной из X-хромосом происходит некое особенное событие, охраняющее ее от воздействия инактивирующего механизма, который тем временем применяется по отношению ко всем остальным X-хромосомам клетки.

Для инактивации X-хромосомы достаточно наличия в ней определенного локуса. Если вдруг между X-хромосомой и аутосомой происходит транслокация и локус попадает в одну из хромосом-химер, этот продукт транслокации будет инактивирован. Локус можно картировать путем сравнения различных транслокаций; называемый центром инактивации X-хромосомы *Xic* (X-inactivation center). Клонированный участок генома протяженностью 450 тпн обладает всеми свойствами *Xic*. Если, в качестве трансгена, такую последовательность внедрить в аутосому, эта аутосома подвергнется инактивации (в системе с культивируемыми клетками).

Содержимое локуса *Xic* оказывает свое действие в *цис*-манере и несет в себе информацию, достаточную для того, чтобы «подсчитать» X-хромосомы и инактивировать их все, кроме одной. Инактивация распространяется от *Xic* по всей X-хромосоме. Если Xic присутствует в хромосоме-химере, возникшей в результате транслокации между X-хромосомой и аутосомой, инактивация распространяется и на аутосомные участки (хотя этот эффект не всегда отличается полнотой).

В *Xic* содержится ген, названный *Xist*, который экспрессируется только на ***неактивной*** X-хромосоме. В плане экспрессии, ген *Xist* ведет себя противоположно остальным генам X-хромосомы, на которой он находится (если *Xist* активен, остальные гены выключены). Делеция *Xist* предотвращает инактивацию X-хромосомы, в которой он отсутствует, но не «сбивает подсчета» X-хромосом (поскольку при этом остальные X-хромосомы могут инактивироваться). Таким образом, у *Xic* можно различить две характерных структурно-функциональных особенности: неустановленный элемент (или элементы), необходимые для подсчета, и ген *Xist*, необходимый для инактивации.

На РИС. 31.12 разъясняется роль РНК *Xist* в инактивации X-хромосомы. Ген *Xist* служит матрицей для синтеза РНК, в которой нет открытых рамок считывания. Вместо того, чтобы кодировать белок, РНК *Xist* «обволакивает» X-хромосому, на которой она была синтезирована. Перед тем, как произойти инактивации, эта РНК синтезируется на обеих X-хромосомах. После инактивации такую РНК дает только неактивная X-хромосома (где ее и можно встретить). Интенсивность же транскрипции *Xist* остается прежней, так что перемена состояния хромосомы зависит не от нее, а от каких-то пост-транскрипционных событий.

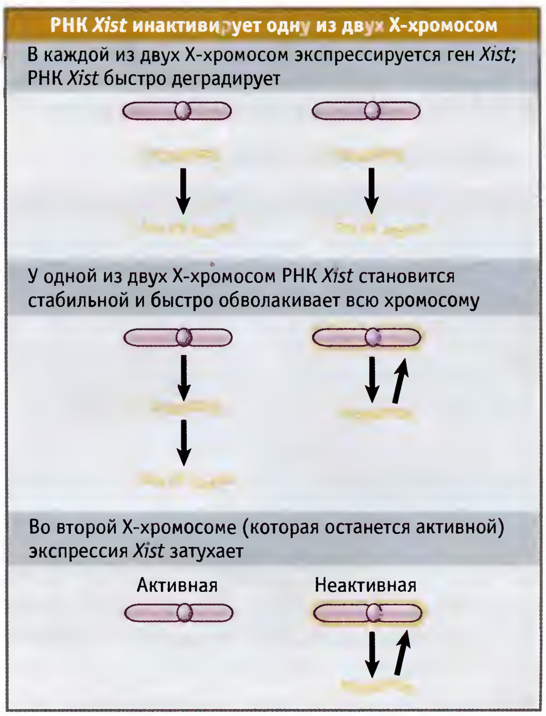


РИС. 31.12 Инактивация X-хромосомы происходит после стабилизации РНК *Xist*, которая в буквальном смысле обволакивает неактивную хромосому

Перед инактивацией X-хромосомы, РНК *Xist* деградирует; период ее полураспада составляет ~2 часа. Инактивация X-хромосомы опосредована стабилизацией РНК *Xist* на неактивной X-хромосоме. РНК *Xist* распределена по длине X-хромосомы в виде точек, указывая на то, что в основу стабилизации может быть заложена ее ассоциация с белками для формирования структуры в виде частиц. Мы пока не знаем, ни какие еще факторы могут быть задействованы в этой реакции, ни почему РНК *Xist* в своем распространении ограничивается своей хромосомой. Такие характерные черты неактивной X-хромосомы, как отсутствие ацетилирования по гистону H4 и метилированность последовательностей CpG (см. разд. 24.19 «CpG-островки являются мишенями для регуляции»), вероятно, обретаются хромосомой позднее, как результат действия механизма инактивации.

Правило n-1 предполагает, что стабилизация РНК *Xist* происходит по умолчанию, и что некий блокирующий механизм предотвращает такую стабилизацию лишь на одной из X-хромосом, которой предстоит остаться активной. Это означает, что хотя наличие Xic необходимо (и достаточно), чтобы хромосома оказалась *инактивирована*, для установления *активного* состояния X-хромосомы могут быть необходимы продукты других локусов.

Для активности X-хромосомы необходимо «погасить» на ней экспрессию *Xist*. Сайленсинг *Xist* можно предотвратить делецией гена ДНК-метилтрансферазы; скорее всего, эффект обусловлен тем, что метилирование промотора *Xist* необходимо для прекращения его транскрипции.

31.6 Конденсация хромосомы — функция конденсинов

Ключевые понятия:

- Белки SMC являются АТФазами; к ним относят конденсины и когезины.
- Гетеродимер белков SMC ассоциирует с другими субъединицами.
- Конденсины сверхспирализуют хроматин за счет того, что вносят в ДНК положительные сверхвитки.
- Конденсины ответственны за конденсацию хромосом при митозе.
- У *C. elegans* особая разновидность конденсинов ответственна за конденсацию X-хромосом, понижающую их активность.

Структурные состояния целых хромосом находятся под влиянием взаимодействий с белками семейства SMC (**SMC, structural maintenance of chromosome**). Эти белки являются АТФазами и подразделяются на две функциональные группы. **Конденсины** участвуют в контроле общей структуры и ответственны за конденсацию хромосом в виде компактных тел при митозе. **Когезины** задействованы в образовании связок между сестринскими хроматидами, которые в ходе

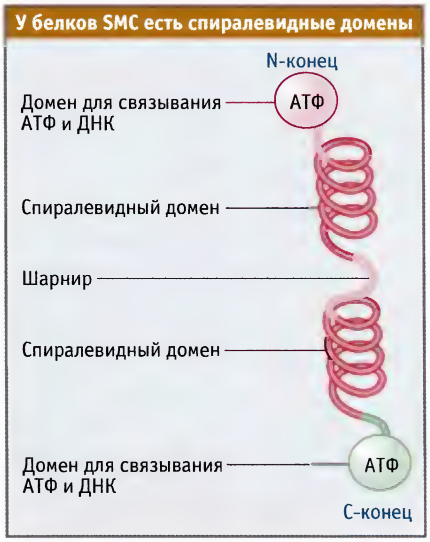


РИС. 31.13 Каждая субъединица SMC — это «модуль Уокера». Концевые компактные домены предназначены для связывания АТФ и ДНК; к ним примыкают спиралевидные домены, разделенные шарниром

митоза должны освободиться друг от друга. И конденсины, и когезины представляют собой димеры белков SMC. Конденсины формируют комплексы, у которых сердцевина, гетеродимер SMC2-SMC4, ассоциирована с другими (не SMC) белками. У когезинов похожая организация строится на сердцевине из гетеродимера SMC1-SMC3.

На РИС. 31.13 показана доменная структура белка SMC. Центральную часть полипептида, структурированную по типу суперспирали, разделяет гибкий шарнирный участок. Обе концевых области полипептида, N и C, содержат АТФ-связывающие и ДНК-связывающие аминокислотные мотивы. Для описания работы белков SMC были предложены разные модели, в зависимости от того, на каких взаимодействиях основывается их димеризация: внутримолекулярных или межмолекулярных.

Эксперименты с бактериальными гомологами белков SMC показали, что их димеры образуются за счет

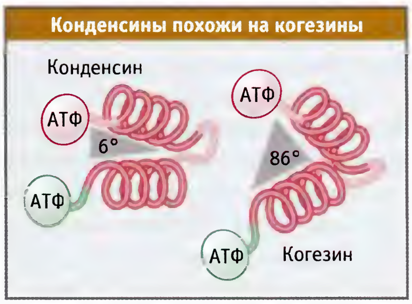


РИС. 31.14 Половинки субъединицы конденсина образуют угол в 6°. Когезин развернут шире, под углом 86°

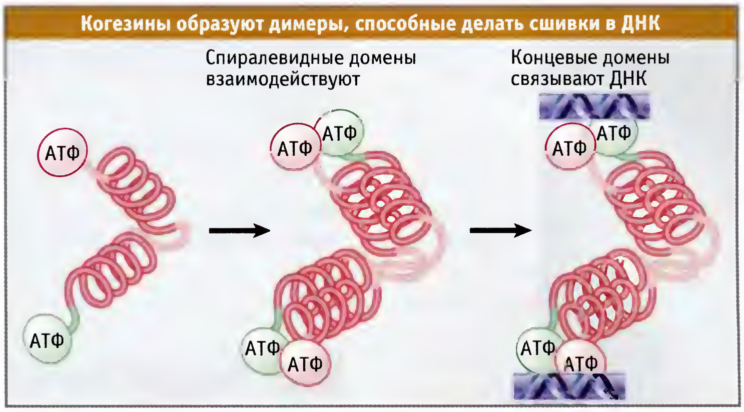


РИС. 31.15 Белки SMC димеризуются за счет антипараллельных взаимодействий между центральными спиралевидными доменами. Оба концевых домена каждой субъединицы SMC содержат аминокислотные мотивы для связывания АТФ и ДНК. Когезины могут образовывать сшивки между двумя разными молекулами ДНК

антипараллельных взаимодействий между суперспиралями. При таком взаимодействии N-конец полипептида одной из субъединиц примыкает к C-концу полипептида другой субъединицы. Благодаря существованию шарнира внутри молекулы SMC, функции когезинов и конденсинов могут определяться разными способами действия со стороны димера. На РИС. 31.14 показано, что когезинам свойственна V-образная структура, с плечами, расходящимися под углом 86°, а в молекулах конденсинов это угол острее, всего 6°. Различие проявляется в том, что когезины удерживают вместе сестринские хроматиды, а конденсины конденсируют отдельную хромосому. На РИС. 31.15 показано, что когезин может принять форму растопыренного димера, сшивающего две молекулы ДНК. На РИС. 31.16 показано, что конденсин может принять форму V-образного димера, в котором основной изгиб приходится на шарнирные участки субъединиц, и который притягивает друг к другу отдаленные участки одной и той же молекулы ДНК, заставляя ее конденсироваться.

Эксперименты с белками дрожжей привели к созданию альтернативной модели, в соответствии с которой белки SMC димеризуются за счет внутримолекулярных взаимодействий, вернее, их гомодимеры образуются за счет одних лишь взаимодействий между идентичными субъединицами. Уже после этого димеры двух разных белков (в конкретном эксперименте, SMC1 и SMC3) могут, взаимодействуя головками и шарнирами, сформировать кольцевую структуру, как это показано на РИС. 31.17. Вместо того, чтобы связать непосредственно ДНК, структура такого типа в состоянии удерживать молекулы ДНК вместе, обхватывая их.

Визуализация митотических хромосом, в сочетании с окрашиванием антителами, демонстрирует, что конденсины распределены по всей длине хромосомы; это видно на РИС. 31.18. (В отличие от конденсинов, когезины распределены вдоль хромосомы по дискретным местоположениям.) Конденсиновый комплекс получил

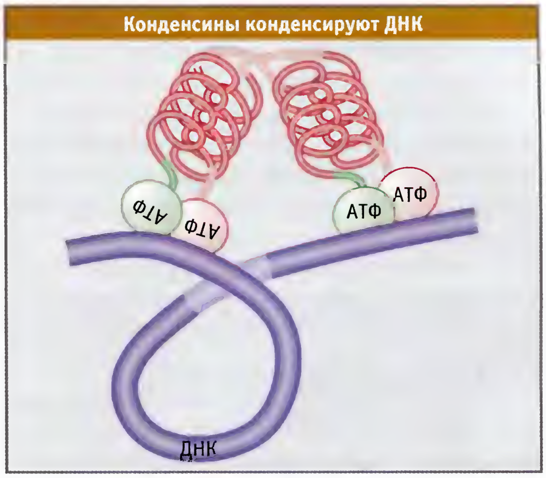


РИС. 31.16 Конденсины, складываясь по шарнирам, могут делать ДНК более компактной

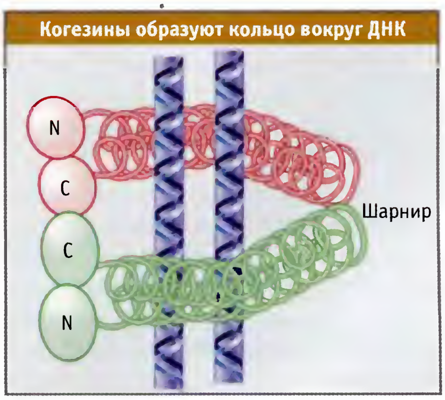


РИС. 31.17 Спиралевидные домены когезинов могут димеризоваться внутри одной субъединицы. В мультимерах субъединицы когезинов контактируют друг с другом концевыми доменами и шарнирами. Окружая молекулы ДНК, мультимеры когезинов могут удерживать их вместе, как пучок



РИС. 31.18 Конденсины разбросаны по всей длине митотической хромосомы. На фотографии красный цвет соответствует ДНК, желтый — конденсинам. Фотография любезно предоставлена Ana Losada and Tatsuya Hirano

свое название за способность заставлять хроматин конденсироваться *in vitro*. Комплекс обладает способностью вносить в ДНК положительные сверхвитки, используя для этого гидролиз АТФ и содействие топоизомеразы I. Эта способность конденсинового комплекса контролируется посредством фосфорилирования его вспомогательных, не SMC, субъединиц; такое фосфорилирование происходит при митозе. Мы пока не знаем, как оно соотнесено с другими модификациями хроматина — например, с фосфорилированием гистонов. Активация конденсинового комплекса специфична для митоза; это заставляет усомниться в том, что комплекс задействован еще и в формировании интерфазного хроматина.

Глобальные изменения хромосом, характеризующиеся изменением статуса их конденсации, связаны с разными типами компенсации доз генов. Так, у самцов *Drosophila* есть специализированный белковый комплекс, локализованный на Х-хромосоме. У *C. elegans* есть белковый комплекс, ассоциированный с обеими Х-хромосомами в клетках эмбрионов XX; компоненты этого комплекса, белки, есть и в эмбрионах XO, но там им свойственно диффузное распределение в клеточном ядре. Комплекс содержит сердцевину из SMC и похож на конденсиновые комплексы, ассоциированные с митотическими хромосомами у животных других видов. Так что комплексу приписывают структурную роль: скорее всего, он заставляет хромосому перейти в более конденсированное, менее активное состояние. Чтобы комплекс распределился по длине Х-хромосомы полностью, могут быть с необходимостью задействованы многочисленные сайты этой хромосомы. Комплекс связывается с этими сайтами, а затем распространяется по хромосоме, покрывая ее более тщательно.

Таким образом, изменения, влияющие на экспрессию всех генов хромосомы отрицательно (как у млекопитающих и *C. elegans*) или положительно (как у *Drosophila*) являются общей чертой механизмов компенсации доз генов. При этом компоненты аппарата компенсации доз генов могут быть различными, точно так же, как варианты их локализации на хромосомах и способы их действия: они в разных случаях различаются.

31.7 Поддерживающая метилаза обеспечивает воспроизводство статуса метилирования ДНК

Ключевые понятия:

- **Большинство метильных групп ДНК обнаружено у цитозинов обеих цепей дублетов CpG.**
- **Репликация переводит полностью метилированный сайт ДНК в два наполовину метилированных сайта.**
- **Полуметилированные сайты превращаются в полностью метилированные сайты под действием фермента — поддерживающей метилазы.**

Метилирование ДНК происходит по специальным сайтам. У бактерий оно ассоциировано с системой рестрикции-метилирования, которой клетка пользуется для защиты от бактериофагов, а также с системой распознавания реплицированной ДНК в сравнении с нереплицированной (см. разд. 20.7 «Контроль направления репарации несовершенных пар оснований»). Из известных функций метилирования ДНК у эукариот, основополагающей является участие в контроле транскрипции: метилирование ассоциировано с инактивацией генов (см. разд. 24.18, «Экспрессия генов связана с деметилированием»).

У животных метилировано от 2% до 7% клеточной ДНК (точный процент зависит от вида животного). Большинство метильных групп обнаружено на «дублетах» CpG, и, на самом деле, метилированными является большинство последовательностей CG. Обычно метилированы цитидиновые группы в обеих цепях короткого палиндрома CG, что придает ему особую структуру.

Сайт ДНК с нестабильными признаками по обеим цепям называют **полностью метилированным**. Рассмотрим последствия репликации такого сайта. На РИС. 31.19 показано, что каждый дочерний дуплекс будет содержать одну метилированную цепь и одну неметилированную цепь. Такой сайт ДНК называют **полуметилированным**.

Теперь воспроизводство метилированного сайта (с сохранением статуса метилирования) зависит от того, что произойдет с полуметилированной ДНК. Случись метилирование неметилированной цепи, и сайт будет восстановлен как полностью метилированный. Если до метилирования произойдет репликация, то полуметилированное состояние воспроизведется лишь в одном из дочерних дуплексов, а другому дочернему

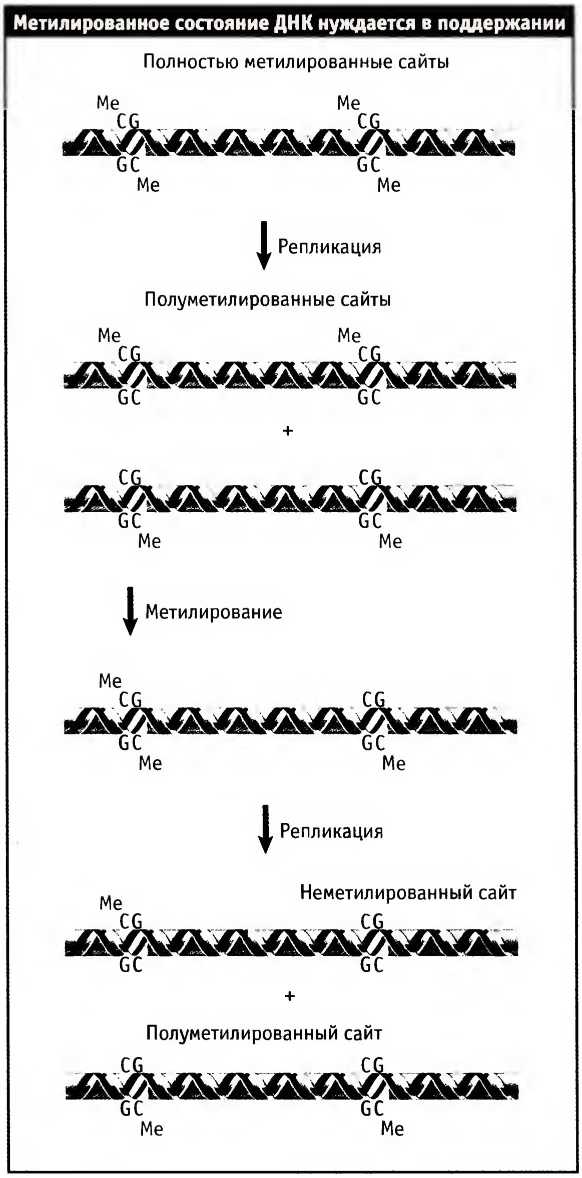


РИС. 31.19 Метилированное состояние определенных сайтов ДНК может поддерживаться ферментом, который узнает в качестве субстрата только полуметилированную ДНК

дуплексу достанется неметилированный сайт. На РИС. 31.20 показано, что статус метилирования ДНК контролируется ферментами, среди которых есть **метилазы** (они добавляют метильные группы по пятому положению цитидина) и **деметилазы** (удаляющие метильные группы). Более формально эти ферменты называют **метилтрансферазами**.

Существует два типа ДНК-метилаз, различающихся по статусу метилирования субстратной ДНК. Чтобы метилировать ДНК «с нуля», необходимо действие ***de novo* метилазы**, узнающей специфическую последовательность ДНК, пригодную для метилирования. Фермент действует *только* на неметилированную ДНК и добавляет метильную группу к одной из ее цепей. У мышей есть две *de novo* метилазы, Dnmt3A и Dnmt3B с разными сайтами-мишенями; обе крайне важны в эмбриональном развитии.

Поддерживающая метилаза конститутивно действует *только на полуметилированные сайты*, превращая их в полностью метилированные сайты. Существование этого фермента означает для клетки воспроизводство каждого метилированного сайта после репликации. У мышей обнаружена одна поддерживающая метилаза, Dnmt1, и она очень важна: мыши с делецией гена метилазы гибнут в раннем эмбриогенезе.

Поддерживающее метилирование эффективно практически на 100%; это гарантирует ситуацию, изображенную на левой стороне рис. 31.19 и превалирующую *in vivo*. В результате, если один из двух, исходно идентичных, аллелей подвергнется метилированию *de novo*, оно будет воспроизводиться в череде клеточных поколений, то есть, возникшее различие между аллелями, никак не обусловленное их последовательностями, будет поддерживаться.

Для метилирования на ДНК существуют разнообразные типы мишеней. Самая распространенная мишень — промоторы генов. Когда ген неактивен, его промотор метилирован, когда же ген работает, его промотор неметилирован. Промоторы мышей, у которых нет Dnmt1, деметилированы. Предполагают, что именно это деметилирование летально, так как оно обуславливает бесконтрольную экспрессию генов. Другой

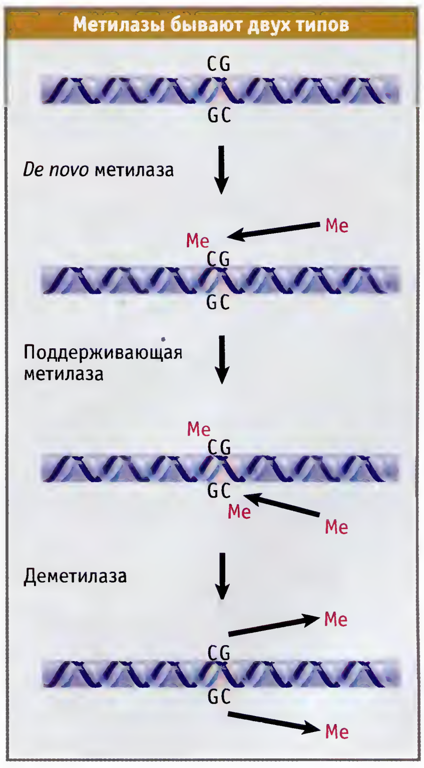


РИС. 31.20 Статус метилирования ДНК контролируется ферментами трех типов. Метилазные активности, *de novo* и поддерживающая, присущи известным белкам; деметилазы пока не идентифицированы

мишенью служит сателлитная ДНК. Мутации по Dnmt3B препятствуют метилированию сателлитной ДНК, что, на клеточном уровне, приводит к нестабильности центромер. Мутации в соответствующем гене человека являются причиной заболевания ICF иммунодефицит/нестабильность центромер, лицевые аномалии. Важность метилирования подчеркивается существованием другого заболевания человека, синдрома Ретта; заболевание вызвано мутацией в гене, кодирующем белок MeCP2, который связывает метилированные последовательности CpG.

В обычном представлении, метилазы — это простые ферменты, самостоятельно действующие на ДНК-мишень. Возможно, наряду с ними существует и другая система метилирования, где для нацеливания фермента на ДНК используется короткая РНК с соответствующей последовательностью (см. разд. 13.6 «Инактивация экспрессии генов с помощью антисмысловой РНК»). Механизм, в соответствии с которым оперирует эта система, пока неясен.

Деметилированные участки ДНК: каким же образом они учреждаются и поддерживаются в таком состоянии? Если сайт ДНК изначально не был метилирован, то белок, узнавший такую (неметилированную) последовательность, может защитить ее от метилирования. Если же какой-либо сайт оказался метилирован, есть два способа его деметилировать. Первый способ — заблокировать работу поддерживающей метилазы на этом сайте во время репликации; тогда, после второго цикла репликации, один из дочерних дуплексов окажется неметилированным (как это показано на правой стороне рис. 31.19). Другой способ заключается в активном деметилировании сайта, как это показано на РИС. 31.21: либо удалив с цитидина метильную группу, либо вырезав метилированный цитозин (или цитидин) из ДНК, предоставив системе репарации заменить его. Известно, что активное деметилирование применяется к отцовской части генома зиготы вскоре после оплодотворения; используемый для этого механизм не известен. Интересно, что в процессе может участвовать цитидин-деаминаза AID; этот фермент умеет деаминировать метилированные цитидины, тем самым создавая в ДНК некомплементарность, которую система репарации может скорректировать, заменив на стандартную (неметилированную) пару G-C.

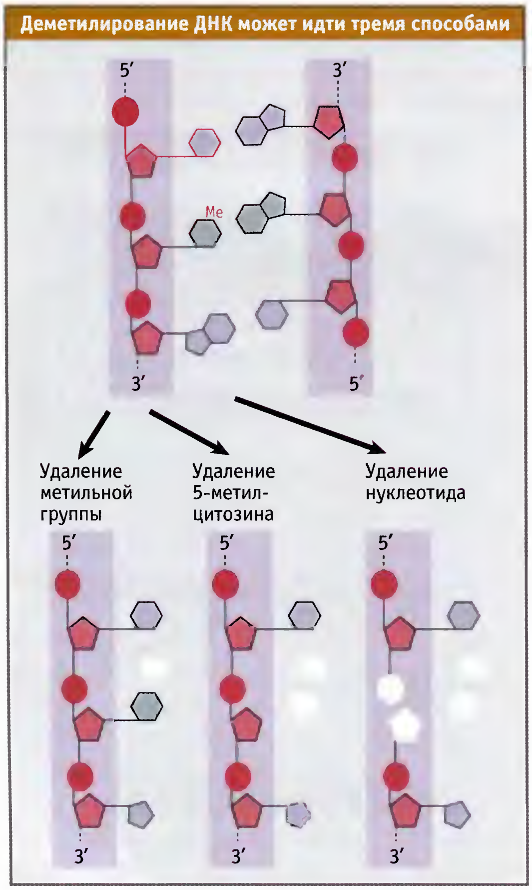


РИС. 31.21 ДНК может быть деметилирована за счет удаления метильной группы, азотистого основания или целого нуклеотида. Отсутствие основания или нуклеотида должно быть восполнено системой репарации

31.8 Метилирование ДНК обусловливает импринтинг

Ключевые понятия:

- При оплодотворении два аллеля одного и того же гена, отцовский и материнский, могут обладать разными паттернами метилирования.
- Метилирование, как правило, ассоциировано с инактивацией гена.
- Если гены метилированы по-разному (импринтированы), эмбрион выживет при условии, что функциональный (не несущий инактивирующей мутации) аллель достался ему от родителя, у которого он не метилирован.
- Выживаемость гетерозигот по импринтированным генам различается в зависимости от направления скрещивания.
- Импринтируемые гены встречаются в составе кластеров, где их статус может находиться под контролем локальной регуляторной области, а метилирование *de novo* происходит в случае отсутствия предотвращающих метилирование механизмов.

Паттерн метилирования ДНК в половых клетках (гаметах) млекопитающих устанавливается во время гаметогенеза. Процесс протекает в две стадии: сначала исходный паттерн оказывается стертым в результате деметилирования в масштабах генома, после чего происходит наложение паттерна, характерного для данного пола.

Когда у эмбриона развиваются примордиальные половые клетки, все различия паттернов метилирования, которые существовали между аллелями до этого момента, теряются. Вне зависимости от пола эмбриона, ранее существовавшие паттерны метилирования оказываются стертыми, а большинство генов — не метилированными. У самцов установление нового паттерна метилирования происходит в два приема. Паттерн, характерный для зре-

лого сперматозоида, устанавливается заранее, на стадии сперматоцита, и последующие изменения вносятся уже после оплодотворения. У самок наложение на ДНК так называемого материнского паттерна метилирования происходит после рождения, когда яйцеклетки, созревая, проходят мейоз.

Как можно предположить, исходя из отсутствия активности генов в гаметах, их типичное состояние — метилированное. Однако по отдельным генам между полами есть различия (когда у одного из полов какой-нибудь локус не метилирован). Возникает вопрос: каким образом детерминируется специфичность метилирования в мужских и женских гаметах?

После оплодотворения, начиная с раннего эмбриогенеза, ДНК зародыша подвержена систематическим изменениям. Некоторые сайты продолжают оставаться метилированными, другие же оказываются специфично не метилированными в клетках, где ген начал экспрессироваться. Судя по паттерну изменений, можно заключить, что индивидуальные события деметилирования, с прицелом на конкретные последовательности ДНК, происходят в период соматического развития, по мере активации конкретных генов.

Специфическим паттерном расположения метильных групп на ДНК (в половых клетках) обусловлен феномен **импринтинга**. Под импринтингом подразумевают разницу, с которой, в плане экспрессии, ведут себя аллели, унаследованные эмбрионом от разных родителей. Так, у эмбрионов мыши экспрессия определенных генов зависит от пола родителя, который передал их эмбриону. Например, аллель, кодирующий инсулиноподобный фактор роста II (IGF-II), унаследованный от отца, экспрессируется, а аллель того же гена, унаследованный от матери, молчит. Это связано с тем, что в ооцитах ген, кодирующий IGF-II, метилирован, а в сперматозоидах — нет; из-за этого в зиготе два аллеля ведут себя по-разному. Такой вариант импринтинга — самый частый, хотя для некоторых генов зависимость активности аллеля от пола родителя, его передавшего, как раз противоположная. Например, противоположный вариант импринтинга, экспрессия материнской копии, был продемонстрирован для гена, кодирующего IGF-IIR (рецептор IGF-II).

Чтобы такой способ наследования работал, необходимо, чтобы паттерн метилирования, с учетом половой специфичности, четко воспроизводился в каждой гамете. На РИС. 31.22 изображена судьба аллелей в гипотетическом локусе мыши. В эмбрионе, на ранней стадии развития, аллель, который был получен от отца, не метилирован и экспрессируется, в то время как материнский аллель метилирован и молчит. Что же происходит, когда у этой мыши начинают формироваться собственные гаметы? Если мышь — самец, то аллель, отходящий в сперматозоид, должен быть не метилированным, безотносительно того, был он исходно метилированным, или нет. Таким образом. Если материнский аллель окажется в сперматозоиде, он, к этому моменту, должен быть деметилирован. Если наша мышь — самка, то аллель отходящий в яйцеклетку, должен быть метилирован; если этот аллель исходно был отцовским, на него должны быть «посажены» метильные группы.

Следствием импринтинга является то, что эмбрион нуждается в отцовском (реже — в материнском) аллеле соответствующего гена. Рассмотрим гетерозиготный эмбрион: у него аллель, пришедший от одного из родителей, несет инактивирующую мутацию, в гомозиготе летальную. Если этот аллель достался эмбриону от отца (а аллель дикого типа унаследован эмбрионом от матери), эмбрион погибнет. Если же аллель дикого типа достался ему, наоборот, от отца, то эмбрион выживет. Такой тип зависимости исхода скрещивания от его направления (генетический термин, означающий важность сочетания пола родителя с вносимой им частью генотипа) не «укладывается» в генетику Менделя. Это — пример эпигенетической наследственности, при которой эффекты генов обуславливает некий фактор, не зависящий от их нуклеотидных последовательностей как таковых (см. разд. 31.10 «Эпигенетические эффекты могут наследоваться»). Несмотря на то, что нуклеотидные последовательности отцовского и материнского аллелей идентичны, эти аллели проявляют разные свойства в зависимости от того, кто из родителей передал их своему потомству. С точки зрения *клеточных* поколений, эти свойства аллелей наследуются через мейоз и последующие серии соматических митозов.

Некоторые импринтируемые (дифференциально метилированные в зависимости от пола родителя) гены собраны в кластеры. Более чем половина из 17 извест-

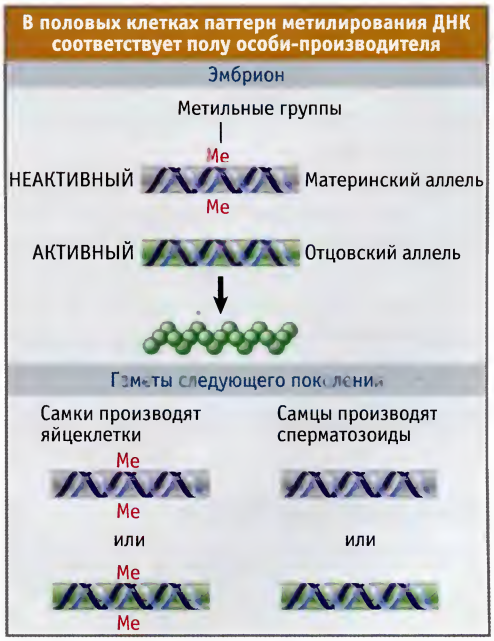


РИС. 31.22 В классическом варианте импринтинга метилированный локус не активен. Предположим, в яйцеклетках какой-нибудь ген метилирован. Тогда у потомства будет активным только отцовский аллель этого гена, и он один будет определять фенотип. В формирующихся гаметах паттерн метилирования ДНК задается в соответствии с полом особи, так что во всех сперматозоидах паттерн метилирования ДНК отцовский, а во всех яйцеклетках — материнский

ных импринтируемых генов мыши сосредоточена в двух конкретных областях генома; в каждой из этих областей есть гены, экспрессируемые у потомства от отца, и гены, экспрессируемые от матери. Это говорит о том, что действие механизмов импринтинга может покрывать в геноме большие расстояния. Об этом также можно судить по делециям в геноме человека, которыми обусловлены синдромы Прадера–Вилли и Ангелмана. В большинстве случаев эти заболевания обусловлены одной и той же делецией четырех мпн ДНК, а различия между двумя синдромами определяются полом родителя, от которого пациенту досталась делеция. Область генома, на которую приходится эта делеция, содержит, по меньшей мере, один ген, импринтируемый от отца и, по меньшей мере, один ген, импринтируемый от матери. Существуют отдельные редкие случаи, когда те же синдромы развиваются на почве гораздо более мелких делеций. Так, синдром Прадера–Вилли может быть вызван делецией 20 тпн ДНК. Эта делеция заставляет замолчать гены, расположенные по обе стороны от нее, причем довольно далеко. Основной эффект делеции состоит в том, что в сперматоцитах мужчин с такой делецией хромосома, унаследованная ими от матери, оказывается не в состоянии обрести паттерн метилирования, требуемый для последующего благополучного отцовства. Это приводит к тому, что гены этой хромосомы передаются потомству в материнском статусе; в результате, молчащими оказываются аллели, полученные как от отца, так и от матери. Эффект того же рода, с «замораживанием», наоборот, отцовского статуса метилирования, обнаруживается при некоторых мелких делециях, вызывающих синдром Ангелмана. Можно предполагать, что участок, на который приходятся эти делеции, заключает в себе некий «центр импринтинга», который, действуя на отдалении, переключает один родительский паттерн метилирования на противоположный.

Также метилированием обусловлены эпигенетичекие эффекты, под контролем которых находится экспрессия рРНК. Эффект **ядрышкового доминирования** заключается в том, что у потомства транскрибируется только один из двух комплектов генов рРНК. Причина эффекта в том, что промоторы генов рРНК, унаследованных от одного из родителей, метилированы по цитозинам.

31.9 Импринтинг генов с противоположной половой специфичностью может контролироваться из одного центра на хромосоме

Ключевые понятия:

- Метилирование может быть причиной инактивации одних генов и активации других генов.
- Импринтинг происходит в результате метилирования *цис*-действующих сайтов.

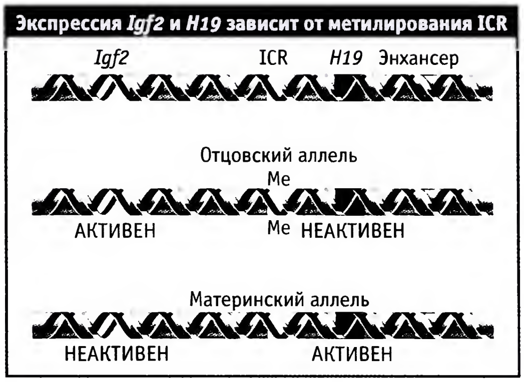


РИС. 31.23 В ДНК, унаследованной от отца, ICR метилирован, из-за этого *Igf2* активен, а *H19* неактивен. В ДНК, унаследованной от матери, ICR, напротив, неметилирован, следовательно, *Igf2* неактивен, *H19* активен

Импринтинг обусловлен статусом метилирования *цис*-действующего сайта поблизости от одного или нескольких генов-мишеней. Участки генома, содержащие такие регуляторные сайты, называют «дифференциально метилированными доменами» (differentially methylated domains, DMD) или «участками контроля импринтинга» (imprinting control regions, ICR). Делеция такого участка нейтрализует импринтинг, после чего гены-мишени в отцовской и материнской частях генома начинают вести себя одинаково.

Насколько разнообразным может быть влияние метилирования ДНК на активность генов, можно судить из наблюдений за областью генома, содержащей два гена, *Igf2* и *H19*. На РИС. 31.23 показано, что эти гены противоположным образом реагируют на статус метилирования ICR, который находится между ними. На отцовской хромосоме ICR метилирован, на что *H19* дает типичный ответ инактивации. Заметим, что *Igf2* при этом экспрес-

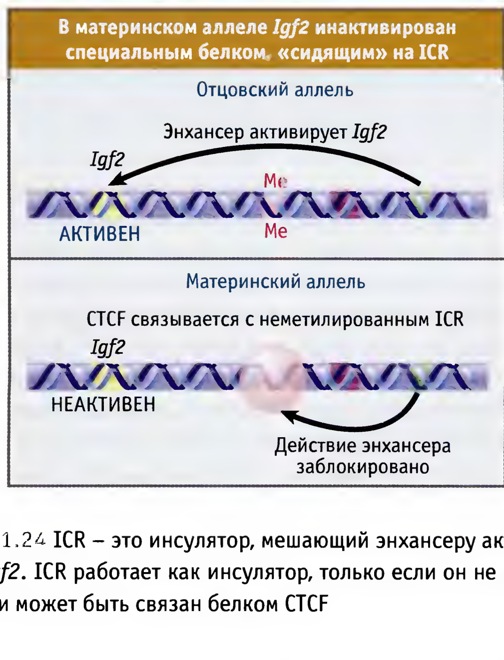


РИС. 31.24 ICR – это инсулятор, мешающий энхансеру активировать *Igf2*. ICR работает как инсулятор, только если он не метилирован и может быть связан белком CTCF

сируется. На материнской хромосоме, где ICR *не*метилирован, имеет место обратная ситуация: *H19* экспрессируется, а *Igf2* инактивирован.

Контроль над *Igf2* зависит от эффективности ICR как инсулятора. На РИС. 31.24 показано, что неметилированный ICR связывает белок CTCF. Это придает ICR свойства инсулятора, запрещающего энхансеру активировать промотор *Igf2*. Эффект необычный: метилирование косвенным образом активирует ген за счет того, что блокирует инсулятор.

В регуляции *H19* использовано более стандартное направление контроля: метилирование переводит ген в неактивное импринтированное состояние, которое может быть прямым следствием ингибирования активности промотора за счет метилирования его ДНК.

31.10 Эпигенетические эффекты могут наследоваться

Ключевые понятия:

- Наследование эпигенетических эффектов может обеспечиваться воспроизводимыми постсинтетическими модификациями нуклеиновых кислот либо за счет самоподдерживающихся белковых структур.

Под **эпигенетическим** наследованием подразумевают способность функциональных состояний генома, у которых могут быть разные фенотипические следствия, наследоваться при отсутствии каких-либо изменений в последовательности ДНК. Как такое может быть? В целом, механизмы возникновения эпигенетических эффектов можно подразделить на два типа:

- ДНК может быть модифицирована химически: к ней, ковалентной связью, прикрепляется группа атомов (например, метильная), и это состояние начинает поддерживаться. В результате два аллеля с одинаковой нуклеотидной последовательностью могут иметь разный статус метилирования, придающий разницу их свойствам.
- Для некоторых белков могут устанавливаться самоподдерживающиеся состояния. Средствами достижения таких состояний могут быть образование комплекса, модификация специфического белка (или белков), принятие белком альтернативной конформации.

Метилирование ДНК обеспечивает эпигенетическое наследование до тех пор, пока конститутивная активность поддерживающей метилазы восстанавливает метилированное состояние после репликации, как показано на рис. 31.19. Статус метилирования ДНК может быть пронесен без изменений через бесконечную череду соматических митозов. Скорее всего, по умолчанию, имеет место именно такая ситуация. Статус метилирования может быть пронесен и через мейоз: например, у гриба *Ascobolus* эпигенетические эффекты могут передаваться дочерним клеткам при митозе и мейозе как раз за счет поддержания статуса метилирования ДНК. У млекопитающих эпигенетические эффекты создаются за счет переустановки статуса метилирования в мужских и женских половых клетках.

Ситуации, при которых эпигенетические эффекты поддерживаются, как будто бы, через состояние белков, изучены, в молекулярном понимании, не так хорошо. Анализ мозаичности, обусловленной эффектом положения, показывает, что конститутивный гетерохроматин может распространиться по геному на некоторое (вариабельное) расстояние, и обретенное состояние будет устойчиво к соматическим делениям. У *Saccharomyces* метилирования ДНК не происходит вовсе, а у *Drosophila* происходит лишь в исчезающе малых дозах, так что наследование эпигенетических состояний, таких как теломерный сайленсинг или та же мозаичность, обусловленная эффектом положения, у этих организмов, по-видимому, полностью зависит от поддержания белковых структур.

На РИС. 31.25 показано, как может сложиться судьба белкового комплекса, ассоциированного с ДНК, после репликации (рассмотрены крайние варианты).

- Комплекс может поддерживать себя, если при репликации разделяется симметрично, и его половинки оказываются ассоциированы с дочерними дуплексами ДНК. Если на основе таких половинок смогут затем сформироваться целые комплексы, исходное состояние будет восстановлено. По смыслу это аналогично поддержанию метилирования

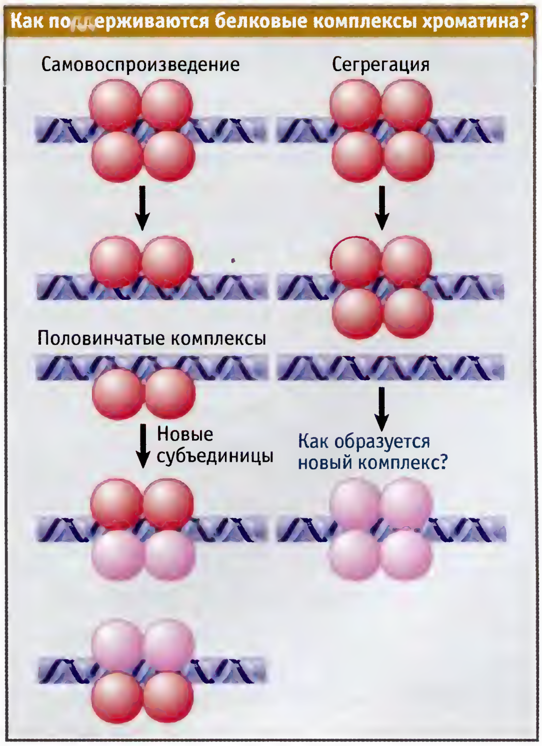


РИС. 31.25 Что происходит с белковыми комплексами хроматина во время репликации?

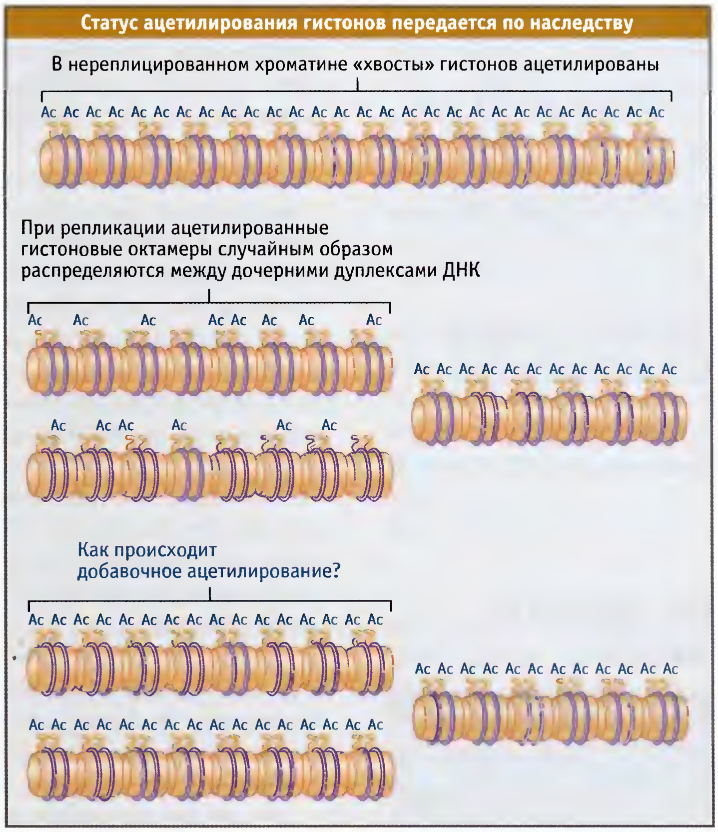


РИС. 31.26 При репликации ацетилированные гистоновые октамеры остаются нетронутыми и распределяются между дочерними волокнами хроматина. Поэтому в каждом дочернем волокне с самого начала есть нуклеосомы, которые имеют «старые» ацетилированные сердцевины

ДНК. Проблема с этой моделью заключается в том, что не очень понятно, зачем белковым комплексам так себя вести.

- Комплекс может сохранить свою целостность и перейти к одному из двух дочерних дуплексов. С этой моделью проблема в том, что в ее рамках новый комплекс (на втором дочернем дуплексе) должен быть собран «с нуля», и не очень понятно, отчего это должно случиться.

Рассмотрим, к примеру, потребность в поддержании структуры гетерохроматина, основанной на присутствии белковых комплексов. Предположим, что белок распределен по отрезку гетерохроматина более или менее непрерывно, как изображено на рис. 31.4. Если при репликации субъединицы белка распределятся по дочерним дуплексам в индивидуальном порядке случайным образом, два дуплекса останутся маркированы белком, несмотря на то, что его плотность понизится вдвое по сравнению с тем, что было до репликации. Если у данного белка есть способность к самосборке, и новые субъединицы стремятся с ним ассоциировать, исходная ситуация может быть восстановлена. *Само существование эпигенетических эффектов заставляет думать, что белок, обусловивший такой эффект, обладает способностью контролировать самовоспроизведение и самосборку.*

В некоторых случаях, эпигенетический эффект может быть обусловлен, скорее, модифицированным состоянием, нежели простым присутствием, белка. Существует общая корреляция между активностью хроматина и статусом ацетилирования гистонов, в особенности N-концевых «хвостов» гистонов H3 и H4. Активация транскрипции сопровождается ацетилированием гистонов вблизи промотора, а репрессия транскрипции ассоциирована с их деацетилированием (см. разд. 30.7 «Гистонацетилазы ассоциированы с активаторами транскрипции»). С наибольшим размахом эта корреляция проявляется в том, что неактивные X-хромосомы в клетках самок млекопитающих «недоацетилированы» по гистону H4.

Для соблюдения инертности конститутивного гетерохроматина может быть необходимо, чтобы его гистоны не были ацетилированы. Если в дрожжевой клетке фермент гистонацетилтрансферазу прикрепить к участку теломерного гетерохроматина, в этом гетерохроматине активируются молчащие гены. Если подвергнуть дрожжи действию трихостатина (ингибитор деацетилирования), их центромерный гетерохроматин станет ацетилированным, и в областях центромер молчащие гены могут стать активными. *Эффект может оставаться в силе даже после снятия обработки трихостатином.* И даже может быть пронесен через митоз

или мейоз. Это говорит о том, что изменение статуса ацетилирования гистонов способно обусловить эпигенетический эффект.

Каким же образом статус ацетилирования гистонов может быть зафиксирован во времени? Предположим, при репликации тетрамеры $H3_2$–$H4_2$ распределяются по двум дочерним дуплексам ДНК случайным образом. Создается ситуация, изображенная на РИС. 31.26: каждый дочерний дуплекс несет октамеры, как полностью ацетилированные по «хвостам» гистонов Н3 и Н4, так и вовсе не ацетилированные. Чтобы эпигенетический эффект закрепился, можно предположить, что наличие на ДНК полностью ацетилированных октамеров служит сигналом к ацетилированию остальных октамеров.

В действительности, ситуация, скорее всего, обстоит сложнее, чем изображено на рисунке, поскольку известно, что во время репликации гистоны, на короткое время, обзаводятся ацетильными группами, поскольку это оптимизирует сборку октамеров. Дальше непонятно: если по завершении монтажа нуклеосом эти ацетильные группы с гистонов (в любом случае) удаляются, то ими можно пренебречь. Альтернативная возможность состоит в том, что полностью ацетилированные октамеры не индуцируют ацетилирование остальных октамеров *de novo*, но, препятствуя деацетилированию гистонов, придают «техническому» ацетилированию регуляторный смысл. А, может быть, они делают и то, и другое.

31.11 Необычное наследование прионов у дрожжей

Ключевые понятия:

- Белок Sup35 в растворимой форме дикого типа является фактором терминации трансляции.
- Тот же белок может существовать в альтернативной форме олигомерного агрегата; эта форма Sup35 не пригодна для участия в синтезе белка.
- Присутствие олигомерной формы Sup35 заставляет новосинтезированный Sup35 принимать конформацию, исключающую его активность.
- Переход Sup35 из одной формы в другую происходит под влиянием шаперонов.
- Форма Sup35 дикого типа имеет рецессивный генетический статус *psi*⁻, а мутантная форма — доминантный генетический статус *PSI*⁺.

Один из самых ясных примеров эпигенетической наследственности, обусловленной состоянием, в котором находится конкретный белок, представляет собой поведение **прионов**. Они представляют собой агрегаты белковых субъединиц с нестандартной конформацией, экспансивные в отношении белковых субъединиц с обычной конформацией. Прионы были обнаружены дважды: в связи с нестандартными наследственными эффектами у дрожжей и в качестве ключевого фактора неврологических заболеваний у млекопитающих, в том числе, у людей. У дрожжей был обнаружен совершенно поразительный эпигенетический эффект: два различных наследственных статуса, [*psi*⁻] и [*PSI*⁺] на генетической карте ложатся в один и тот же локус, причем *нуклеотидная последовательность гена в этом локусе в обоих случаях одна и та же*. Переключение соответствующих фенотипов с одного на другой в дрожжевых клетках случается редко и выглядит как результат самопроизвольной смены наследственного статуса.

Генотип *psi*, по генетической карте, соответствует локусу *SUP35*. Там находится одноименный ген, кодирующий фактор терминации трансляции. На РИС. 31.27 показаны все возможные эффекты присутствия белка Sup35 в дрожжевой клетке. В клетках «дикого типа», характеризуемых как [*psi*⁻], ген активен, и его белок, растворимый фактор Sup35, терминирует синтез белков на рибосомах. В клетках-мутантах [*PSI*⁺] фактор Sup35 не функционирует, что приводит к неспособности терминировать синтез белка надлежащим образом. (Эта особенность была исходно зафиксирована на основании летальных последствий улучшения эффективности супрессоров охра-кодонов в штаммах [*PSI*⁺].)

С точки зрения генетики, штаммы [*PSI*⁺] обладают необычными свойствами. Когда штамм [*psi*⁻] скрещивают со штаммом [*PSI*⁺], *все потомство получается* [*PSI*⁺]. У дрожжей, такого паттерна наследования можно было бы

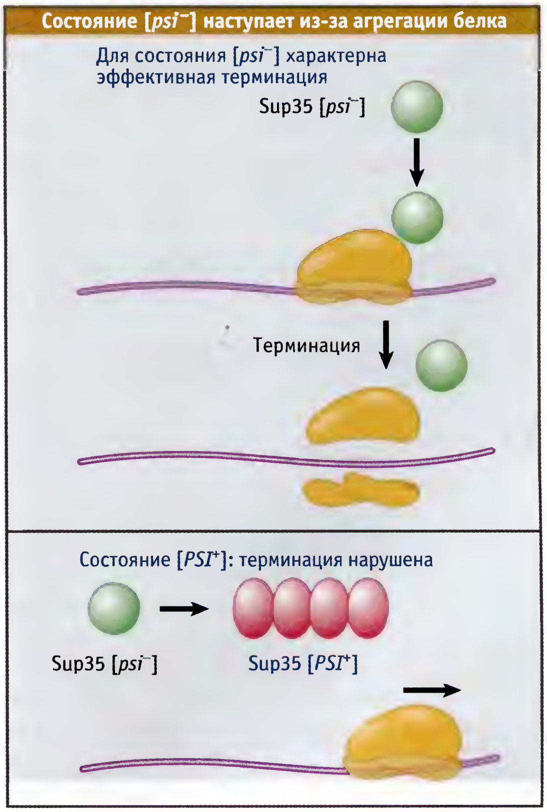


РИС. 31.27 Эффективность терминации трансляции зависит от конформации белка Sup35

ожидать от внехромосомного элемента, однако фенотип [*PSI*⁺] не привязан ни к одной нуклеиновой кислоте такого рода. Фенотип [*PSI*⁺] метастабилен; этот термин означает, что, хотя фенотип наследуется подавляющим большинством потомства, его потеря происходит чаще, чем можно было бы ожидать вследствие одних только мутаций.

Фенотип клеток [*PSI*⁺] обусловлен конформацией белка Sup35. В клетке дикого типа, [*psi*⁻], этот белок нормальным образом выполняет свою функцию. В клетке [*PSI*⁺] белок присутствует в альтернативной конформации, с потерей своей нормальной функции. Чтобы объяснить одностороннее доминирование *PSI*⁺ над *psi*⁻ при скрещиваниях, придется предположить, что ***присутствие белка*** Sup35 ***в состоянии, характерном для [PSI⁺], заставляет остальной клеточный белок*** Sup35 ***перейти в такое же состояние.*** Для этого потребовалось бы взаимодействие между Sup35[*PSI*⁺] и новосинтезированным Sup35. Такое взаимодействие может представлять собой наработку олигомерной формы белка, в которой Sup35[*PSI*⁺] принадлежит роль затравки, как показано на **РИС. 31.28**.

Кроме *SUP35*, похожим образом проявляет себя локус *URE2*; в нем закодирован белок, необходимый для

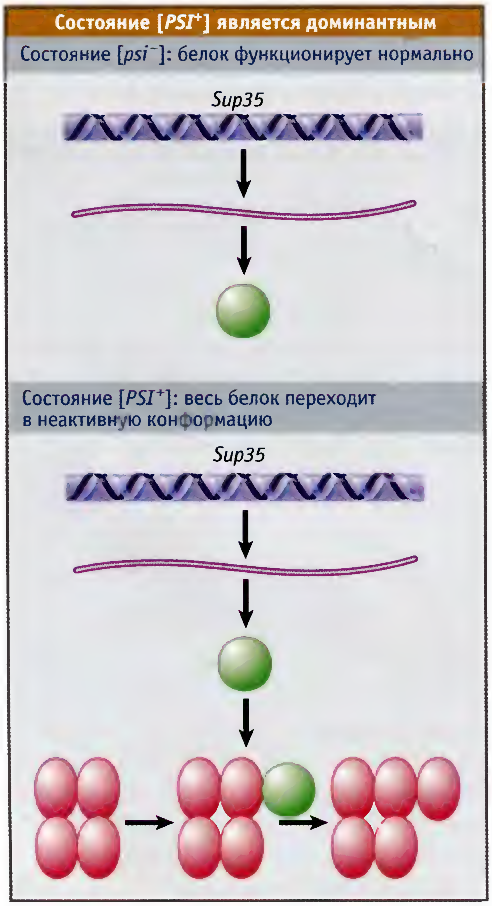


РИС. 31.28 Если в клетке есть прион, характерный для состояния [*PSI*⁺], новосинтезированный белок Sup35 переходит в неактивную конформацию

азот-зависимой репрессии некоторых катаболических ферментов. При переходе штамма в альтернативное состояние [*URE3*] белок Ure2 прекращает функционировать (при этом штамм обретает генетический статус *URE3*). У белков Sup35 и Ure2 есть сходство: каждый состоит из двух доменов, функционирующих независимо. Наличие С-концевого домена является достаточным условием активности белка. Наличия N-концевого домена достаточно для формирования структур, которые делают белок неактивным. Поэтому дрожжи, у которых N-концевой домен Sup35 отсутствует из-за делеции, не могут стать [*PSI*⁺], а присутствия в клетке N-концевого домена Sup35 в конформации, характерной для [*PSI*⁺], достаточно, чтобы удерживать весь клеточный Sup35 в форме Sup35[*PSI*⁺]. Принципиальной особенностью N-концевого домена является обилие глютаминов (Gln) и аспарагинов (Asn).

Потеря функции Sup35 в клетках [*PSI*⁺] происходит из-за перевода всего клеточного пула Sup35 в состав олигомерных комплексов. В клетках [*PSI*⁺] белок Sup35 скапливается в дискретных центрах (фокусах), а в клетках [*psi*⁻] тот же белок диффузно рассеян в цитоплазме. *In vitro*, белок Sup35 из клеток [*PSI*⁺] образует амилоидные фибриллы с характерным высоким содержанием β-слоев.

Значимость именно конформации белка (а не ковалентных модификаций) можно выявить, создав определенные условия. Так, мягкая денатурирующая обработка приводит к выходу клетки из состояния [*PSI*⁺].

Наследственный характер состояния [*PSI*⁺] обеспечивает, в частности, шаперон Hsp104. Манера его действия довольно парадоксальна. Делеция Hsp104 препятствует поддержанию состояния [*PSI*⁺], но и перегрузка клетки белком Hsp104 также приводит к потере состояния [*PSI*⁺]. Возможно, Hsp104 необходим для какого-то изменения в структуре Sup35, необходимого для обретения конформации Sup35[*PSI*⁺], причем это изменение обязательно должно носить промежуточный характер.

На способности Sup35 формировать неактивную структуру *in vitro* строится биохимическое доказательство ключевой роли этого белка в смене клеточных состояний. **РИС. 31.29** иллюстрирует поразительный эксперимент: белок, переведенный в неактивную форму *in vitro*, был помещен в липосомы (по сути, окружен искусственной мембраной), а затем внесен непосредственно в клетки, посредством слияния липосом с клеточными мембранами дрожжей [*psi*⁻]. Дрожжи перешли в состояние [*PSI*⁺]! Этот результат отметает все возражения, которым подвергались умозаключения о том, что белок способен задавать эпигенетически-наследуемое состояние. В экспериментах, где клетки проходят спаривание, или где одни клетки обрабатывают экстрактами других клеток, всегда существует опасность случайного «несанкционированного» переноса нуклеиновой кислоты. В только что описанном эксперименте ситуация другая: разница между препаратом белка, который вызывает переход клеток в другое состояние, и препаратом, который такого перехода не вызывает, заключается только в обработке белка, а ее проводили *in vitro*; таким образом,

возможность засора нуклеиновой кислотой была полностью исключена.

Способность дрожжей переходить в состояние [PSI^+], обусловленное прионами, зависит от благоприятного сочетания других факторов наследственности (генетического фона). Чтобы переход в [PSI^+] состоялся, дрожжи должны быть [PIN^+]. Состояние [PIN^+] само по себе также является эпигенетическим. Оно может возникнуть за счет образования прионов на основе любого из нескольких разных белков. Подобно Sup35, эти белки имеют Gln/Asn-богатые домены. Сверхэкспрессия этих доменов у дрожжей стимулирует переход в состояние [PSI^+].

Каким же образом присутствие одного Gln/Asn-богатого белка может влиять на образование прионов совсем другим белком? Мы знаем, что формирование прионов Sup35 специфично для белка Sup35, то есть, никакого смешения (перекрестной агрегации) с другими белками при этом не происходит. Возникает предположение, что дрожжевая клетка может содержать растворимые белки, противодействующие образованию прионов. В результате, повышение концентрации любого Gln/Asn-богатого белка, который будет с ними взаимодействовать, снизит их концентрацию, что позволит другим белкам с Gln/Asn-богатыми доменами агрегировать беспрепятственно.

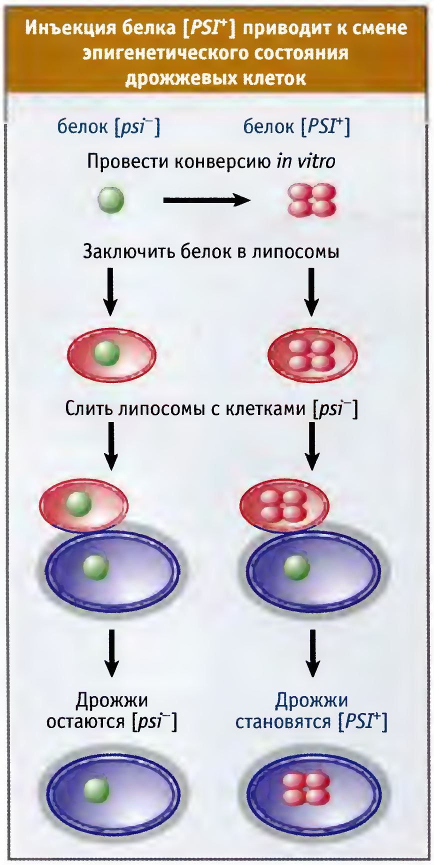


РИС. 31.29 Под действием очищенного белка [PSI^+] происходит конверсия дрожжей из состояния [psi^-] в состояние[PSI^+]

31.12 Прионы как причина заболеваний млекопитающих

Ключевые понятия:

- Белок, которым обусловлено заболевание скрэпи у овец, существует в двух формах; неинфекционная форма дикого типа, PrP^C, расщепляется протеазами, а болезнетворная форма, PrP^{Sc}, устойчива к протеазам.
- Неврологическое заболевание может быть передано от овец мышам при помощи инъекции очищенного белка, приона PrP^{Sc}.
- Чтобы заболеть, мышь-реципиент должна иметь функциональную копию гена *PrP*.
- Прион PrP^{Sc} может поддерживать себя, заставляя новосинтезированный белок PrP принимать форму PrP^{Sc} вместо PrP^C.
- Прион PrP^{Sc} существует в разных вариантах, отличающихся между собой конформацией субъединиц.

Болезни, вызываемые прионами, обнаружены у овец, коров и людей. Основной симптом, атаксия, представляет собой нейродегенеративное расстройство, проявляющееся в неспособности сохранять прямое положение тела. Название болезни у овец — почесуха, или **скрэпи** — имеет отношение к ее внешнему проявлению: чтобы стоять прямо, овца трется о стены. Здоровую овцу можно заразить скрэпи, привив ей экстракт ткани больного животного. У папуасов Новой Гвинеи была обнаружена болезнь **куру**, которая, по-видимому, передавалась за счет каннибализма (в особенности, поедания мозгов). У западных народов есть похожие заболевания: например, синдром Герстманна–Штрауслера, предрасположенность к которому передается по наследству, и похожую болезнь Кройцфельда–Якоба (CJD), которая возникает спорадически. Не так давно была описана болезнь, напоминающая CJD и вызванная, судя по всему, потреблением мяса коров, страдавших коровьим бешенством.

Если ткань овцы, больной скрэпи, привить мыши, то через 75–150 дней мышь заболеет. Активный инфекционный агент представляет собой белок, устойчивый к протеазам. Этот белок закодирован геном, экспрессируемым в мозге. Форму белка, присутствующую в нормальном мозге, называют PrP^C; она чувствительна к протеазам (то есть, расщепляется протеазами). Переход белка в другую форму, PrP^{Sc}, устойчивую к протеазам, ассоциирован с заболеванием. Инфекционный препарат, приготовленный на основе тканей больного животного, не содержит нуклеиновых кислот и чувствителен к УФ-облучению на длинах волн, при которых повреждается белок. Препарат обладает низкой инфективностью (1 инфекционная единица на 10^5 белковых субъединиц PrP^{Sc}). Такое описание подпадает под категорию эпигенетической наследственности, при которой, в отсутствие изменения генетической информации (здоровые клетки и больные клетки несут одну и ту же последовательность гена *PrP*), форма PrP^{Sc} белка PrP является инфекционным агентом, в то время как другая его форма, PrP^C, безвредна. Форма PrP^{Sc} отличается высоким содержанием β-слоев, которые способствуют образованию амилоидной

фибриллярной структуры; форма PrP^C такую структуру не образует.

Основа различий между формами PrP^{Sc} и PrP^C, по всей видимости, состоит в изменении конформации *без* помощи каких-либо ковалентных модификаций. Обе формы белка гликозилированы и закреплены в мембране глицерол-фосфатидил-инозитольным якорем (GPI-якорь).

На РИС. 31.30 изображены результаты некоторых экспериментов. В обычной ситуации, белок PrP^{Sc}, экстрагированный из больной мыши и привитый здоровой мыши, индуцирует у нее развитие смертельного заболевания. Если у здоровой мыши ген *PrP* «нокаутирован», такая мышь, получив инъекцию PrP^{Sc}, ***не заболеет.*** Здесь важно обратить внимание на два момента. Во-первых, эндогенный белок необходим для инфекции, потому что он обеспечивает сырье для наработки новых количеств инфекционного агента. Во-вторых, причина болезни не в устранении PrP^C; болезнь вызывается за счет обретения белком новой функции при переходе в форму PrP^{Sc}. Интересно, что если ген *PrP* изменен, и белок, который он кодирует, не может быть заякорен в мембране при помощи GPI, то у мышей, гомозиготных по такому измененному гену, болезнь не развивается.

Экспериментальная система с заражением мышей подходит для изучения зависимости инфекционной активности белка от его аминокислотной последовательности. Кроме того, факт развития данного заболевания у животных разных видов служит основой для изучения межвидовых барьеров, возникающих в его передаче от животного к животному, а также позволяет конструировать гибридные белки и по их совокупности определять общие черты белковой структуры, которыми может быть обусловлена инфективность PrP. Препаратами скрэпи, полученными от овец, удалось заразить (и вызвать развитие заболевания) животных нескольких видов. Но уже между этими видами животных заболевание передается не всегда. Например, мыши устойчивы к инфекции прионами хомячка. По сути, это означает, что PrP^{Sc} хомячка не может перевести PrP мыши из формы PrP^C в форму PrP^{Sc}. Изменить ситуацию с заражением мышей прионами хомячка можно, если заменить у мыши ее собственный ген PrP на ген PrP хомячка. Для этого надо ввести ген PrP хомячка в мышь с «нокаутированным» PrP (вывести такую новую линию мышей, трансгенных по PrP хомячка, на основе «нокаутной» линии PrP^-/PrP^-). Мышь с геном PrP хомячка восприимчива к инфекции белком PrP^{Sc} хомячка. Это говорит о том, что для конверсии клеточного белка PrP^C в форму PrP^{Sc} требуется определенная степень структурного сходства белков.

У мышей существует множество разновидностей PrP^{Sc}; их можно распознать по характерной продолжительности инкубационного периода, от инокуляции (прививки) до появления симптомов. Это означает, что белок не ограничивается двумя альтернативными формами, PrP^C и PrP^{Sc}. Скорее всего, есть некоторое множество состояний PrP^{Sc}. Различия между ними должны выражаться в способности к экспансии, но не в аминокислотной последовательности (которая у всех форм одинаковая). Согласно такой концепции, если PrP^{Sc} отличается от PrP^C

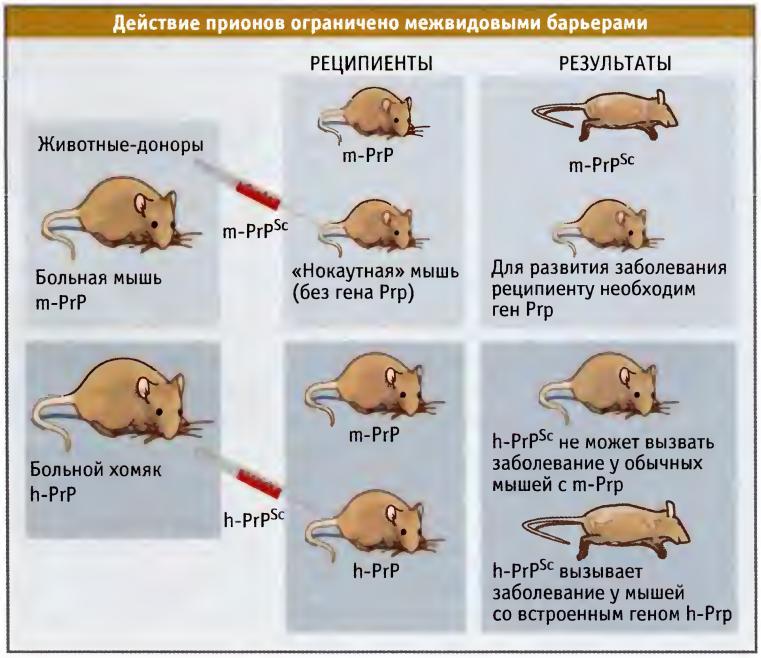


РИС. 31.30 Инъекция белка Prp^{Sc} приводит к заболеванию в том случае, если инъецированный Prp^{Sc} и эндогенный PrP^C животного относятся к одному типу

конформацией, то этих альтернативных патогенных конформаций может быть много, причем каждая из них способна переводить в себе подобную конформацию нормальный PrP^{C}».

Вероятность перехода PrP^{C} в форму PrP^{Sc} во многом зависит от аминокислотной последовательности PrP. Так, развитие синдрома Герстманна–Штраусслера у людей обусловлено одной-единственной аминокислотной заменой в PrP; предрасположенность к заболеванию наследуется как доминантный признак. Если такую же замену произвести у мышей (изменив для этого нуклеотидную последовательность гена *PrP*), у них развивается аналогичная болезнь. По всей видимости, мутантный белок подвержен повышенному риску перехода в форму PrP^{Sc}. Похожим образом, предрасположенность овец к спонтанному заболеванию скрэпи может быть оценена по их генотипам, а именно, по комбинации аминокислот в трех конкретных положениях. Комбинация проверяется анализом нуклеотидного полиморфизма по кодонам 136, 154 и 171 гена *PrP*.

Прионы воплощают собой экстремальный случай эпигенетической наследственности: инфекционным агентом является белок, способный принимать множество конформаций, каждая из которых способна к самовоспроизводству. Это свойство позволяет переводить идентифицированный белок в ту же конформацию, что и у приона.

31.13 Резюме

Формирование гетерохроматина происходит благодаря особым белкам: одни из них связываются со специализированными областями хромосом (например, с теломерами), другие взаимодействуют с гистонами. Стартовав в центре инициации, неактивная структура может, в своем формировании, распространиться по хроматиновой нити. У дрожжей похожие события инактивации имеют место в локусах, содержащих молчащие кассеты для детерминации типа спаривания. У *Drosophila* репрессивные структуры, необходимые для поддержания определенных генов в неактивном состоянии, обеспечиваются комплексом, в который входят белки Pc-G. Эти структуры своей способностью распространяться от инициирующего центра напоминают гетерохроматин.

Формирование гетерохроматина может быть инициировано на определенных участках генома, после чего оно может распространиться по геному на расстояние, заранее точно не предопределенное. После перехода некоторой области генома в состояние гетерохроматина, это состояние будет пронесено данной областью через последующие деления; таким образом, состояние гетерохроматина, в которое перешла фиксированная область генома, будет наследоваться. В результате возникает паттерн эпигенетической наследственности, согласно ее принципу, две абсолютно одинаковых последовательности ДНК могут быть ассоциированы с разными белковыми структурами и, в силу этого, отличаться по своей способности к экспрессии. Так, у *Drosophila* мозаицизм обусловлен эффектом положения и возникает из-за вариабельности в распространении эффекта состояния гетерохроматина по геномной последовательности, в сочетании с последующим четким эпигенетическим наследованием границ гетерохроматина в клеточных линиях.

Модификация «хвостов» в молекулах гистонов работает на запуск реорганизации хроматина; ацетилирование, в общем случае, ассоциировано с активацией генов. Гистонацетилазы можно встретить в активирующих комплексах, а гистондеацетилазы входят в состав репрессирующих (инактивирующих) комплексов. Метилирование гистонов ассоциировано с инактивацией генов. Некоторые модификации гистонов по отношению к другим модификациям, могут быть взаимоисключающими, другие — синергичными.

У дрожжей репрессия активности хроматина в теломерах и локусах, где хранятся молчащие кассеты для детерминации типа спаривания, происходит схожим образом: за счет взаимодействия определенных белков с N-концевыми «хвостами» гистонов H3 и H4. Образование инактивирующего комплекса может быть инициировано связыванием одного из белков со специальной последовательностью ДНК; после этого другие компоненты комплекса кооперативно полимеризуются вдоль геномного окружения этой последовательности.

Инактивация одной из X-хромосом у самок плацентарных млекопитающих в эмбриональных клетках происходит случайным образом. Локус *Xic* необходим и достаточен для «подсчета» клеткой числа своих X-хромосом. Правило n–1 гласит, что в каждой клетке инактивируются все X-хромосомы кроме одной. Локус *Xic* содержит ген *Xist*; его транскрипт, РНК *Xist*, не кодирует белок. После события инактивации РНК *Xist* присутствует только на неактивной X-хромосоме. Стабилизация РНК *Xist* является ключевым фактором становления инактивированного состояния X-хромосомы.

Статус метилирования клеточной ДНК эпигенетически наследуется клеточным потомством. Непосредственные продукты репликации ДНК являются полуметилированными; активность поддерживающей метилазы доводит их до полностью метилированного состояния. Статус метилирования некоторых сайтов ДНК в хромосомах млекопитающих может различаться в зависимости от того, какой из родителей передал своему потомку данную хромосому. Геномные ДНК в сперматозоидах и яйцеклетках обладают специфическими и отличными друг от друга паттернами метилирования. В результате аллели, доставшиеся от отца, и аллели, доставшиеся от матери, экспрессируются в клетках эмбриона по-разному. С этим обстоятельством связано явление импринтинга. Классический вариант импринтинга заключается в том, что неметилированный аллель, унаследованный от одного из родителей, существен для эмбриона, потому

что только этот аллель активен; другой аллель, унаследованный от родителя противоположного пола, молчит из-за того, что ДНК метилирована. Паттерны метилирования, специфичные для гамет конкретного пола, устанавливаются на ДНК при образовании гамет у особей данного пола, в каждом поколении.

Прионы — это инфекционные агенты белковой природы, которыми обусловлено развитие заболевания скрэпи у овец и похожих заболеваний у людей. Инфекционный агент представляет собой разновидность нормального клеточного белка. Эта разновидность (PrP^{Sc}) имеет измененную конформацию и обладает способностью к самовоспроизводству путем преобразования нормальных PrP^{c}-белков. Нормальный белок, PrP^{C}, не принимает такую конформацию в обычных условиях, но переходит в нее в присутствии PrP^{Sc}. Наследование элемента [*PSI*] у дрожжей обстоит похожим образом.

Литература

31.2 Гетерохроматин распространяется от места зарождения

Исследование

Ahmad, K. and Henikoff, S. (2001). Modulation of a transcription factor counteracts heterochromatic gene silencing in *Drosophila. Cell* 104, 839–847.

31.3 Существование гетерохроматина обусловлено взаимодействиями гистонов с другими белками

Обзор

Loo, S. and Rine, J. (1995). Silencing and heritable domains of gene expression. *Annu. Rev. Cell Dev. Biol.* 11, 519–548.
Moazed, D. (2001). Common themes in mechanisms of gene silencing. *Mol. Cell* 8, 489–498.
Rusche, L. N., Kirchmaier, A. L., and Rine, J. (2003). The establishment, inheritance, and function of silenced chromatin in *Saccharomyces cerevisiae. Annu. Rev. Biochem.* 72, 481–516.
Thompson, J. S., Hecht, A., and Grunstein, M. (1993). Histones and the regulation of het-erochromatin in yeast. *Cold Spring Harbor Symp. Quant. Biol.* 58, 247–256.
Zhang, Y. and Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev.* 15, 2343–2360.

Исследование

Ahmad, K. and Henikoff, S. (2001). Modulation of a transcription factor counteracts heterochromatic gene silencing in *Drosophila. Cell* 104, 839–847.
Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410, 120–124.
Bloom, K. S. and Carbon, J. (1982). Yeast centromere DNA is in a unique and highly ordered structure in chromosomes and small circular minichromosomes. *Cell* 29, 305–317.
Cheutin, T., McNairn, A. J., Jenuwein, T., Gilbert, D. M., Singh, P. B., and Misteli, T. (2003). Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science* 299, 721–725.
Eissenberg, J. C., Morris, G. D., Reuter, G., and Hartnett, T. (1992). The heterochromatin-associated protein HP-1 is an essential protein in *Drosophila* with dosage-dependent effects on position-effect variegation. *Genetics* 131, 345–352.
Hecht, A., Laroche, T., Strahl-Bolsinger, S., Gasser, S. M., and Grunstein, M. (1995). Histone H3 and H4 N-termini interact with the silent information regulators SIR3 and SIR4: a molecular model for the formation of heterochromatin in yeast. *Cell* 80, 583–592.
Imai, S., Armstrong, C. M., Kaeberlein, M., and Guarente, L. (2000). Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. *Nature* 403, 795–800.
James, T. C. and Elgin, S. C. (1986). Identification of a nonhistone chromosomal protein associated with heterochromatin in *D. melanogaster* and its gene. *Mol. Cell Biol.* 6, 3862–3872.
Kayne, P. S., Kim, U. J., Han, M., Mullen, R. J., Yoshizaki, F., and Grunstein, M. (1988). Extremely conserved histone H4 N terminus is dispensable for growth but essential for repressing the silent mating loci in yeast. *Cell* 55, 27–39.
Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410, 116–120.
Landry, J., Sutton, A., Tafrov, S. T., Heller, R. C., Stebbins, J., Pillus, L., and Sternglanz, R. (2000). The silencing protein SIR2 and its homologs are NAD-dependent protein deacetylases. *Proc. Natl. Acad. Sci. USA* 97, 5807–5811.
Manis, J. P., Gu, Y., Lansford, R., Sonoda, E., Ferrini, R., Davidson, L., Rajewsky, K., and Alt, F. W. (1998). Ku70 is required for late B cell development and immunoglobulin heavy chain class switching. *J. Exp. Med.* 187, 2081–2089.
Meluh, P. B. et al. (1998). Cse4p is a component of the core centromere of *S. cerevisiae. Cell* 94, 607–613.
Moretti, P., Freeman, K., Coodly, L., and Shore, D. (1994). Evidence that a complex of SIR proteins interacts with the silencer and telomere-binding protein RAP1. *Genes Dev.* 8, 2257–2269.
Nakagawa, H., Lee, J. K., Hurwitz, J., Allshire, R. C., Nakayama, J., Grewal, S. I., Tanaka, K., and Murakami, Y. (2002). Fission yeast CENP-B homologs nucleate centromeric het-erochromatin by promoting heterochromatin-specific histone tail modifications. *Genes Dev.* 16, 1766–1778.
Nakayama, J., Rice, J. C., Strahl, B. D., Allis, C. D., and Grewal, S. I. (2001). Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* 292, 110–113.
Palladino, F., Laroche, T., Gilson, E., Axelrod, A., Pillus, L., and Gasser, S. M. (1993). SIR3 and SIR4 proteins are required for the positioning and integrity of yeast telomeres. *Cell* 75, 543–555.
Platero, J. S., Hartnett, T., and Eissenberg, J. C. (1995). Functional analysis of the chromo domain of HP1. *EMBO J.* 14, 3977–3986.
Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R., and Reuter, G. (2002). Central role of *Drosophila* SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. *EMBO J.* 21, 1121–1131.
Sekinger, E. A. and Gross, D. S. (2001). Silenced chromatin is permissive to activator binding and PIC recruitment. *Cell* 105, 403–414.
Shore, D. and Nasmyth, K. (1987). Purification and cloning of a DNA-binding protein from yeast that binds to both silencer and activator elements. *Cell* 51, 721–732.
Smith, J. S., Brachmann, C. B., Celic, I., Kenna, M. A., Muhammad, S., Starai, V. J., Avalos, J. L., Escalante-Semerena, J. C., Grubmeyer, C., Wolberger, C., and Boeke, J. D. (2000). A phylogenetically conserved NAD+-dependent protein deacetylase activity in the Sir2 protein family. *Proc. Natl. Acad. Sci. USA* 97, 6658–6663.
Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S. I., and Moazed, D. (2004). PHKi-mediated targeting of heterochromatin by the RITS complex. *Science* 303, 672–676.

31.4 Репрессоры и активаторы: белки-антагонисты Polycomb и Trithorax

Обзор

Ringrose. L. and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu. Rev. Genet.* 38, 413–443.

Исследование

Brown, J. L., Fritsch, C., Mueller, J., and Kassis, J. A. (2003). The *Drosophila* pho-like gene encodes a YY1-related DNA binding protein that is redundant with pleiohomeotic in homeotic gene silencing. *Development* 130, 285–294.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* 298, 1039–1043.

Chan, C. S., Rastelli, L., and Pirrotta, V. (1994). A Polycomb response element in the Ubx gene that determines an epigenetically inherited state of repression. *EMBO J.* 13, 2553–2564.

Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A., and Pirrotta, V. (2002). *Drosophila* enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* 111, 185–196.

Eissenberg, J. C., James, T. C., Fister-Hartnett, D. M., Hartnett, T., Ngan, V., and Elgin, S. C. R. (1990). Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in *D. melanogaster. Proc. Natl. Acad. Sci. USA* 87, 9923–9927.

Fischle, W., Wang, Y., Jacobs, S. A., Kim, Y., Allis, C. D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromo domains. *Genes Dev.* 17, 1870–1881.

Francis, N. J., Kingston, R. E., and Woodcock, C. L. (2004). Chromatin compaction by a Polycomb group protein complex. *Science* 306, 1574–1577.

Franke, A., DeCamillis, M., Zink, D., Cheng, N., Brock, H. W., and Paro, R. (1992). Polycomb and polyhomeotic are constituents of a multimeric protein complex in chromatin of *D. melanogaster. EMBO J.* 11, 2941–2950.

Geyer, P. K. and Corces, V. G. (1992). DNA position-specific repression of transcription by a *Drosophila* zinc finger protein. *Genes Dev.* 6, 1865–1873.

Orlando, V. and Paro, R. (1993). Mapping Polycomb-repressed domains in the bithorax complex using *in vivo* formaldehyde cross-linked chromatin. *Cell* 75, 1187–1198.

Strutt, H., Cavalli, G., and Paro, R. (1997). Colocalization of Polycomb protein and GAGA factor on regulatory elements responsible for the maintenance of homeotic gene expression. *EMBO J.* 16, 3621–3632.

Wang, L., Brown, J. L., Cao, R., Zhang, Y., Kassis, J. A., and Jones, R. S. (2004). Hierarchical recruitment of Polycomb group silencing complexes. *Mol. Cell* 14, 637–646.

Zink, B. and Paro, R. (1989). *In vivo* binding patterns of a trans-regulator of the homeotic genes in *D. melanogaster. Nature* 337, 468–471.

31.5 Х-хромосомы претерпевают глобальные изменения

Обзор

Plath, K., Mlynarczyk-Evans, S., Nusinow, D. A., and Panning, B. (2002). Xist PHK and the mechanism of X chromosome inactivation. *Annu. Rev. Genet.* 36, 233–278.

Исследование

Jeppesen, P. and Turner, B. M. (1993). The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell* 74, 281–289.

Lee, J. T. et al. (1996). A 450 kb transgene displays properties of the mammalian X-inactivation center. *Cell* 86, 83–94.

Lyon, M. F. (1961). Gene action in the X chromosome of the mouse. *Nature* 190, 372–373.

Panning, B., Dausman, J., and Jaenisch, R. (1997). X chromosome inactivation is mediated by Xist PHK stabilization. *Cell* 90, 907–916.

Penny, G. D. et al. (1996). Requirement for Xist in X chromosome inactivation. *Nature* 379, 131–137.

31.6 Конденсация хромосомы — функция конденсинов

Обзор

Hirano, T. (1999). SMC-mediated chromosome mechanics: a conserved scheme from bacteria to vertebrates? *Genes Dev.* 13, 11–19.

Hirano, T. (2000). Chromosome cohesion, condensation, and separation. *Annu. Rev. Biochem.* 69, 115–144.

Hirano, T. (2002). The ABCs of SMC proteins: two-armed ATPases for chromosome condensation, cohesion, and repair. *Genes Dev.* 16, 399–414.

Jessberger, R. (2002). The many functions of SMC proteins in chromosome dynamics. *Nat. Rev. Mol. Cell Biol.* 3, 767–778.

Nasmyth, K. (2002). Segregating sister genomes: the molecular biology of chromosome separation. *Science* 297, 559–565.

Исследование

Csankovszki, G., McDonel, P., and Meyer, B. J. (2004). Recruitment and spreading of the *C. elegans* dosage compensation complex along X chromosomes. *Science* 303, 1182–1185.

Haering, C. H., Lowe, J., Hochwage, A., and Nasmyth, K. (2002). Molecular architecture of SMC proteins and the yeast cohesin complex. *Mol. Cell* 9, 773–788.

Kimura, K., Rybenkov, V. V., Crisona, N. J., Hirano, T., and Cozzarelli, N. R. (1999). 13S condensin actively reconfigures DNA by introducing global positive writhe: implications for chromosome condensation. *Cell* 98, 239–248.

31.7 Поддерживающая метилаза обеспечивает воспроизводство статуса метилирования ДНК

Обзор

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev.* 16, 6–21.

Bird, A. P. (1986). A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Nature* 321, 209–213.

Matzke, M., Matzke, A. J., and Kooter, J. M. (2001). PHK: guiding gene silencing. *Science* 293, 1080–1083.

Sharp, P. A. (2001). PHK interference—2001. *Genes Dev.* 15, 485–490.

Исследование

Amir, R. E., Van den Veyver, I. B., Wan, M., Tran, C. Q., Francke, U., and Zoghbi, H. Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.* 23, 185–188.

Li, E., Bestor, T. H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69, 915–926.

Morgan, H. D., et al. (2004). Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J. Biol.Chem.* 279, 52353–52360.

Okano, M., Bell, D. W., Haber, D. A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99, 247–257.

Xu, G. L., Bestor, T. H., Bourc'his, D., Hsieh, C. L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J. J., and Viegas-Paquignot, E. (1999). Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* 402, 187–191.

31.8 Метилирование ДНК обусловливает импринтинг

Обзор

Bartolomei, M. S. and Tilghman, S. (1997). Genomic imprinting in mammals. *Annu. Rev. Genet.* 31, 493–525.

Исследование

Chaillet, J. R., Vogt, T. F., Beier, D. R., and Leder, P. (1991). Parental-specific methylation of an imprinted transgene is established during gametogenesis and progressively changes during embryogenesis. *Cell* 66, 77–83.

Lawrence, R. J., Earley, K., Pontes, O., Silva, M., Chen, Z. J., Neves, N., Viegas, W., and Pikaard, C. S. (2004). A concerted DNA methylation/histone methylation switch regulates rPHK gene dosage control and nucleolar dominance. *Mol. Cell* 13, 599–609.

31.9 Импринтинг генов с противоположной половой специфичностью может контролироваться из одного центра на хромосоме

Исследование

Bell, A. C. and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* 405, 482–485.

Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M., and Tilghman, S. M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 405, 486–489.

31.11 Необычное наследование прионов у дрожжей

Обзор

Horwich, A. L. and Weissman, J. S. (1997). Deadly conformations: protein misfolding in prion disease. *Cell* 89, 499–510.

Lindquist, S. (1997). Mad cows meet psi-chotic yeast: the expansion of the prion hypothesis. *Cell* 89, 495–498.

Serio, T. R. and Lindquist, S. L. (1999). [PSI$^+$]: an epigenetic modulator of translation termination efficiency. *Annu. Rev. Cell Dev. Biol.* 15, 661–703.

Wickner, R. B. (1996). Prions and PHK viruses of *S. cerevisiae*. *Annu. Rev. Genet.* 30, 109–139.

Wickner, R. B., Edskes, H. K., Roberts, B. T., Baxa, U., Pierce, M. M., Ross, E. D., and Brachmann, A. (2004). Prions: proteins as genes and infectious entities. *Genes Dev.* 18, 470–485.

Wickner, R. B., Edskes, H. K., Ross, E. D., Pierce, M. M., Baxa, U., Brachmann, A., and Shewmaker, F. (2004). Prion genetics: new rules for a new kind of gene. *Annu. Rev. Genet.* 38, 681–707.

Исследование

Chernoff, Y. O. et al. (1995). Role of the chaperone protein Hsp104 in propagation of the yeast prion-like factor [PSI$^+$]. *Science* 268, 880–884.

Derkatch, I. L., Bradley, M. E., Hong, J. Y., and Liebman, S. W. (2001). Prions affect the appearance of other prions: the story of [PIN(+)]. *Cell* 106, 171–182.

Derkatch, I. L., Bradley, M. E., Masse, S. V., Zadorsky, S. P., Polozkov, G. V., Inge-Vechtomov, S. G., Liebman S. W. (2000). Dependence and independence of [PSI(+)] and [PIN(+)]: a two-prion system in yeast? *EMBO J.* 19, 1942–1952.

Glover, J. R. et al. (1997). Self-seeded fibers formed by Sup35, the protein determinant of [PSI+], a heritable prion-like factor of *S. cerevisiae*. *Cell* 89, 811–819.

Masison, D. C. and Wickner, R. B. (1995). Prion-inducing domain of yeast Ure2p and protease resistance of Ure2p in prion-containing cells. *Science* 270, 93–95.

Osherovich, L. Z. and Weissman, J. S. (2001). Multiple gln/asn-rich prion domains confer susceptibility to induction of the yeast. *Cell* 106, 183–194.

Sparrer, H. E., Santoso, A., Szoka, F. C., and Weissman, J. S. (2000). Evidence for the prion hypothesis: induction of the yeast [PSI+] factor by in vitro-converted Sup35 protein. *Science* 289, 595–599.

Wickner, R. B. (1994). [URE3] as an altered URE2 protein: evidence for a prion analog in *S. cerevisiae*. *Science* 264, 566–569.

31.12 Прионы как причина заболеваний млекопитающих

Обзор

Chien, P., Weissman, J. S., and DePace, A. H. (2004). Emerging principles of conformation-based prion inheritance. *Annu. Rev. Biochem.* 73, 617–656.

Prusiner, S. (1982). Novel proteinaceous infectious particles cause scrapie. *Science* 216, 136–144.

Prusiner, S. B. and Scott, M. R. (1997). Genetics of prions. *Annu. Rev. Genet.* 31, 139–175.

Wickner, R. B., Edskes, H. K., Roberts, B. T., Baxa, U., Pierce, M. M., Ross, E. D., and Brachmann, A. (2004). Prions: proteins as genes and infectious entities. *Genes Dev.* 18, 470–485.

Исследование

Basler, K., Oesch, B., Scott, M., Westaway, D., Walchli, M., Groth, D. F., McKinley, M. P., Prusiner, S. B., and Weissmann, C. (1986). Scrapie and cellular PrP isoforms are encoded by the same chromosomal gene. *Cell* 46, 417–428.

Bueler, H. et al. (1993). Mice devoid of PrP are resistant to scrapie. *Cell* 73, 1339–1347.

Hsiao, K. et al. (1989). Linkage of a prion protein missense variant to Gerstmann-Straussler syndrome. *Nature* 338, 342–345.

McKinley, M. P., Bolton, D. C., and Prusiner, S. B. (1983). A protease-resistant protein is a structural component of the scrapie prion. *Cell* 35, 57–62.

Oesch, B. et al. (1985). A cellular gene encodes scrapie PrP27-30 protein. *Cell* 40, 735–746.

Scott, M. et al. (1993). Propagation of prions with artificial properties in transgenic mice expressing chimeric PrP genes. *Cell* 73, 979–988.

Словарь терминов

10-нм фибрилла. Цепочки нуклеосом, образующиеся в результате декомпактизации хроматина.

30-нм фибрилла. Спираль из нуклеосом. Базовый уровень организации нуклеосом в составе хроматина.

45S РНК. Предшественник, включающий в себя последовательность обеих больших рибосомальных РНК (28S и 18S рРНК).

5,8S РНК. Независимая малая РНК, входящая в состав большой субъединицы эукариотической рибосомы. Гомологична 5'-концевой части бактериальной 23S рРНК.

5S РНК. 120-нуклеотидная РНК, входящая в состав большой субъединицы рибосомы.

Alu-домен. Включает в себя участки 7S РНК SRP-частицы, гомологичные Alu-последовательностям.

ARS. Дрожжевой ориджин репликации. Общей чертой различных последовательностей *ARS* является наличие консервативной последовательности из 11 пн, называемой А-доменом.

***att*-сайты.** Участки фага лямбда и бактериальной хромосомы, по которым фаг интегрирует или вырезается из бактериальной хромосомы за счет рекомбинации.

Bam-островки. Серии коротких повторяющихся последовательностей, обнаруженные в нетранскрибируемых спейсерах генов рРНК шпорцевой лягушки. Название отражает их обнаружение при помощи эндонуклеазы рестрикции BamI.

bZIP. Белок, обладающий основным ДНК-связывающим регионом, соединенным с димеризующим мотивом типа «лейциновая молния».

B-клетка. Лимфоцит, продуцирующий антитела. Образование B-клеток идет преимущественно в костном мозге.

B-клеточная память. Способность B-клеток к быстрой продукции антител в ходе повторного иммунного ответа. B-клетки памяти продуцируют антитела более высокой аффинности, чем наивные B-клетки.

B-клеточный рецептор. Рецептор к антигену на поверхности B-лимфоцита, состоящий из мембраносвязанного иммуноглобулина, нековалентно присоединенного к Igα- и Igβ-цепям.

B-форма ДНК. Правозакрученная двойная спираль с десятью парами нуклеотидов на виток (360°) спирали. Открытая Криком и Уотсоном, наиболее распространенная при нормальных условиях форма ДНК.

CAAT-бокс. Часть консервативной последовательности, расположенной перед стартовыми точками эукариотических транскрипционных единиц; узнается обширной группой факторов транскрипции.

CD3. Белковый комплекс, который ассоциирует с цепями α и β рецептора к антигену на поверхности T-клетки. Каждый комплекс включает в себя по одной цепи δ, ε и γ и две цепи ζ.

CpG-островок. Участок генома млекопитающего, богатый неметилированными дублетами CpG. Такие участки имеют протяженность 1-2 тпн.

C-концевой домен (CTD) эукариотической РНК-полимеразы II. Подлежит фосфорилированию во время инициации. Он участвует в координировании нескольких энзиматических активностей, связанных с транскрипцией.

C-полосы. Яркие темные точки — внешний вид центромер на препаратах хромосом, окрашенных с использованием специальных реактивов, взаимодействующих с центромерами.

D-петля. Участок митохондриальной ДНК, где одна из ее цепей взаимодействует с коротким участком РНК, вытесняющей комплементарную цепь из дуплекса с первой. Тот же самый термин используется для описания вытеснения участка одной из цепей дуплекса за счет инвазии идентичной одноцепочечной молекулы.

D-сегмент. Дополнительная последовательность, находящаяся между участками V и J тяжелой цепи иммуноглобулина.

D-шпилька тРНК. Отличается высоким содержанием азотистого основания дигидроуридина.

EF-Tu. Фактор элонгации, который связывает аминоацил-тРНК и помещает ее в сайт A бактериальной рибосомы.

GC-бокс. Элемент, общий для промоторов ДНК-полимеразы II; представляет собой последовательность GGGCGG.

GMP-PCP. Химический аналог ГТФ, который не может быть гидролизован. Используется в исследованиях сложных реакций — для вычленения стадий, требующих гидролиза ГТФ.

G-полосы. Поперечная исчерченность эукариотических хромосом, наблюдаемая на препаратах, окрашенных по специальной методике. Паттерны G-полос используются для кариотипирования (идентификации конкретных хромосом и их конкретных областей).

Hb анти-Лепор. Фьюжн-ген, продукт неравного кроссинговера. Его белок представляет собой химеру N-концевой области β-глобина с C-концевой областью δ-глобина.

Hb Кения. Фьюжн-ген, продукт неравного кроссинговера между генами глобинов Aγ и β.

Hb Лепор. Фьюжн-ген, продукт неравного кроссинговера между генами глобинов β и γ. Его белок, в целом похожий на β-гло-

бин, представляет собой химеру N-концевой области δ-глобина с С-концевой областью β-глобина.

Hydrops fetalis. Смертельное заболевание, вызванное отсутствием гена гемоглобина α.

IF-1. Бактериальный фактор инициации, который стабилизирует комплекс инициации.

IF-2. Бактериальный фактор инициации, который связывает инициаторную тРНК с комплексом инициации.

IF-3. Бактериальный фактор инициации, нужный для связывания 30S субъединицы с сайтами инициации в мРНК. Он также мешает связыванию 30S субъединицы с 50S субъединицей.

Inr. Участок промотора РНК-полимеразы II между позициями −3 и +5; обобщенная последовательность читается как Py_2CAPy_5. Это простейший из возможных промоторов pol II.

J-сегменты. Кодирующие последовательности в локусах иммуноглобулинов и рецепторов Т-клеток. J-сегменты расположены между вариабельными (V) и константными (C) сегментами этих генов.

N-нуклеотидная последовательность. Короткая последовательность, синтезированная без матрицы, которая случайным образом добавляется в места кодирующих соединений во время перестройки гена иммуноглобулина или гена Т-клеточного рецептора. N-нуклеотиды увеличивают разнообразие рецепторов к антигенам.

N-формил-метионил-тРНК ($тРНК_f^{Met}$). Аминоацил-тРНК, инициирующая синтез белка у бактерий. Аминогруппа метионина формилирована.

PrP. Белок, активный компонент приона, вызывающего почесуху овец (скрэпи) и ряд похожих заболеваний. Форма PrP, приводящая к развитию заболевания, называется PrP^{Sc}.

P-нуклеотидная последовательность. Короткий палиндром, вырезаемый из гена при перестановке сегментов V-, (D)-, J-генных сегментов иммуноглобулинов и Т-клеточных рецепторов. P-нуклеотиды отделяются от кодирующих соединений, поскольку белки RAG обрезают концевые шпильки, образующиеся во время перестройки.

P-элемент. Разновидность транспозонов *D. melanogaster*.

rec-мутации *E. coli* приводят к невозможности гомологичной рекомбинации.

RF1. Бактериальный релизинг-фактор, который узнает триплеты UAA и UAG в качестве сигналов терминации синтеза белка.

RF2. Бактериальный релизинг-фактор, который узнает триплеты UAA и UGA в качестве сигналов терминации синтеза белка.

RF3. Бактериальный фактор терминации синтеза белка, похожий на фактор элонгации EF-G. Его функция заключается в освобождении фактора RF1 или RF2 из рибосомы.

***rut*.** Последовательность РНК, узнаваемая фактором терминации rho.

R-сегменты. Повторяющиеся последовательности на краях ретровирусной РНК. Обозначаются как R-U5 и U3-R.

S фаза. Ограниченная по времени часть эукариотического клеточного цикла, в течение которой происходит синтез ДНК.

SL-РНК. Небольшая РНК, отдающая экзон в реакции транс-сплайсинга у трипаносом и нематод.

SOS-ответ. Скоординированная индукция множества ферментов в клетке *E. coli* (включая системы репарации), в ответ на радиацию или другие множественные повреждения ДНК. Является следствием активации протеолитической активности RecA, разрушающей репрессор LexA.

SOS-сайт. Последовательность ДНК (оператор) длиной около 20 пн, узнаваемая репрессором LexA.

SR-белок. Участвующий в сплайсинге белок, обладающий вариабельным по длине участком с большим количеством аргининов и серинов.

SWI/SNF. Комплекс перестройки хроматина, использующий гидролиз АТФ для изменения организации нуклеосом.

S-домен. Последовательность 7S РНК в SRP; пространственно отделена от Alu-домена.

S-участок. Последовательность, задействованная в переключении класса иммуноглобулинов; состоит из повторяющихся последовательностей на 5′-концах сегментов, кодирующих константные участки тяжелых цепей.

TAF. Субъединицы транскрипционного фактора $TF_{II}D$, помогающие TBP связаться с ДНК. Также они создают точки взаимодействия с другими компонентами аппарата транскрипции.

$TF_{II}D$. Фактор транскрипции, который связывает последовательность TATA, которая предшествует точке начала транскрипции в промоторах для РНК-полимеразы II. $TF_{II}D$ состоит из TBP и субъединиц TAF, которые связывают TBP.

TATA-бокс. Консервативный A-T-богатый октамер, обнаруживаемый примерно в 25 пн перед точкой начала транскрипции эукариотической РНК-полимеразой II. Последовательность вовлечена в позиционирование фермента на промоторе для корректной инициации синтеза РНК.

TATA-связывающий белок (TBP). Субъединица транскрипционного фактора $TF_{II}D$, связывающая ДНК.

Tn. Обозначение бактериальных транспозонов, несущих маркеры, не относящиеся к их функции (например, гены устойчивости к антибиотикам); разновидности обозначают номерами (например, Tn10).

***Ty*-транспозон.** Первый мобильный элемент, обнаруженный в дрожжах (назван от transposon yeast).

Т-хелпер. Т-лимфоцит, который активирует макрофагов, стимулирует пролиферацию В-лимфоцитов и продукцию антител. Т-хелперы обычно экспрессируют поверхностный CD4, но не CD8.

U3. Повторяющаяся последовательность на 3′-конце ретровирусной РНК.

U5. Повторяющаяся последовательность на 5′-конце ретровирусной РНК.

V-ген. Последовательность, кодирующая значительную часть вариабельной (N-концевой) области иммуноглобулина.

Абортивная инициация. Процесс, в ходе которого РНК-полимераза начинает транскрипцию, прерывающуюся после синтеза нескольких первых нуклеотидов, после чего еще не покинувшая промотор РНК-полимераза начинает все заново. Может пройти несколько таких циклов перед тем, как синтез РНК перейдет в стадию элонгации.

Авирулентные мутанты. Бактерии или вирусы, утратившие способность к продуктивной инфекции.

Автономный контролирующий элемент кукурузы. Активный транспозон, обладающий способностью к транспозиции (см. неавтономный контролирующий элемент).

Автосплайсинг. Способность интрона к самовырезанию из РНК путем каталитической реакции, зависящей только от последовательности данного интрона.

Агропиновые плазмиды. Плазмиды, несущие гены синтеза опинов агропиновой группы.

Адаптивный иммунитет. Альтернативное название приобретенного иммунитета.

Аденилатциклаза. Фермент, использующий АТФ в качестве субстрата для синтеза циклоАМФ (в котором 5′- и 3′-позиции сахара соединены фосфатной группой).

А-домен. Консервативная последовательность из 11 пар А-Т в составе дрожжевого ARS-элемента (ориджина репликации).

Акридины. Мутагены, действие которых на ДНК вызывает вставки или делеции отдельных нуклеотидов. Активно использовались при выяснении триплетной структуры генетического кода.

Аксиальный элемент. Белковая структура, вокруг которой хромосомы конденсируются при образовании синапсиса.

Активатор CRP. Белок, позитивный регулятор транскрипции, активируемый циклоАМФ. Необходим для инициации транскрипции во многих оперонах *E. coli.*

Активатор. Белок, усиливающий экспрессию гена путем связывания промотора и привлечения к нему РНК-полимеразы. У эукариот участок связывания активатора в составе промотора называют регуляторным элементом.

Акцепторный конец тРНК. Короткий дуплекс в составе тРНК, заканчивающийся последовательностью ССА, к которой прикрепляется аминокислота.

Аллель. Одна из двух (или нескольких) альтернативных форм гена.

Аллельное исключение. Экспрессия в лимфоците только одного аллеля, кодирующего иммуноглобулин. Определяется наличием механизма блокирования активации аллеля на другой хромосоме при запуске одного из аллелей.

Аллостерическая регуляция. Способность фермента менять конформацию (и, как следствие, активность) при связывании небольшой молекулы в специальном сайте, отличном от субстрат-связывающего сайта и активного центра фермента.

Альтернативный сплайсинг. Образование нескольких разных РНК-продуктов из одного предшественника за счет использования при сплайсинге разных экзон-интронных стыков.

Аманитин (α-аманитин). Бицикличный октапептид из ядовитого гриба *Amanita phalloides*, ингибирующий работу некоторых эукариотических РНК-полимераз, в особенности РНК-полимеразы II.

Амбер-кодон. Триплет UAG, один из трех кодонов, обуславливающих терминацию белкового синтеза.

Аминоацил-тРНК-синтетазы. Ферменты, осуществляющие ковалентное присоединение аминокислот к 2′- или 3′-ОН-концам тРНК.

Аминоацил-тРНК. тРНК, несущая аминокислоту. СООН-группа аминокислоты соединяется с 2′- или 3′-ОН-группой концевого нуклеотида тРНК.

Анти-Sm. Аутоиммунная сыворотка, обусловливающая Sm-эпитоп, общий с группой белков, обнаруженных в Sm-РНП, вовлеченных в сплайсинг РНК.

Антиген трансплантации. Белок, кодируемый локусом главного комплекса гистосовместимости, присутствующий во всех клетках млекопитающих и участвующий во взаимодействии между лимфоцитами.

Антиген. Молекула, способная специфично связываться с антигеновым рецептором (таким, как антитело).

Антиинсулятор. Последовательность, позволяющая энхансеру преодолеть эффект инсулятора.

Антикодон. Триплет, занимающий определенное и постоянное положение в структуре молекулы тРНК; комплементарно взаимодействует с кодоном (или кодонами) мРНК.

Антикодоновая петля. Структура типа «стебель-петля» в составе транспортной РНК, несущая антикодон.

Антипараллельные цепи двойной спирали. Цепи ДНК, организованные в противоположной ориентации, так что 5′-конец одной цепи взаимодействует с 3′-концом другой цепи.

Антисмысловая цепь ДНК. Цепь ДНК, комплементарная смысловой цепи, выступающая в качестве матрицы для синтеза мРНК.

Антисмысловой ген. Ген, кодирующий антисмысловую РНК, которая комплементарна мРНК-мишени и блокирует ее функцию.

Антитело. Белок, продуцируемый В-лимфоцитами и связывающий конкретный антиген. Антитела синтезируются в мембрансвязанной и секретируемой формах. Антитела продуцируются в ходе иммунного ответа для нейтрализации и удаления патогенов.

Антитерминаторные белки. Белки, позволяющие РНК-полимеразе игнорировать некоторые терминаторы.

Антитерминация. Механизм контроля транскрипции, при котором терминация предотвращается на специальном сайте, позволяя РНК-полимеразе считать гены, расположенные за ним.

А-сайт. Сайт посадки аминоацил-тРНК на рибосому и взаимодействия с кодоном мРНК.

Аск. Орган гриба, содержащий четыре или восемь гаплоидных спор, представляющих собой продукт одного мейоза.

Ассимиляция одиночной цепи. Способность белка RecA замещать одноцепочечной ДНК гомологичную цепь в дуплексе, приводя к вхождению одночепочечнй ДНК в состав дуплекса.

Аттенуатор. Терминатор, участвующий в аттенуации.

Аттенуация. Способ регуляции бактериальных оперонов путем контроля терминации транскрипции по сайтам, расположенным перед первым структурным геном.

Аутогенный контроль. Ингибирующее (негативный аутогенный контроль) или активирующее (позитивный аутогенный контроль) действие продукта гена на его собственную экспрессию.

Аутоиммунное заболевание. Патологическое состояние, при котором иммунный ответ направлен на собственный антиген.

Ацентрический фрагмент хромосомы. Возникший в результате двуцепочечного разрыва фрагмент хромосомы, теряемый в ходе деления из-за отсутствия центромеры.

Базовый фактор. Транскрипционный фактор, необходимый РНК-полимеразе II для формирования комплекса инициации на любых промоторах. Факторы обозначают как $TF_{II}X$, где X — буква.

Безнуклеоидная бактерия. Бактерия, потерявшая нуклеоид, но сохранившая схожую с диким типом форму.

Белки MHC класса I. Одна из преобладающих разновидностей молекул MHC. В большинстве случаев белки MHC класса I осуществляют презентацию пептидов цитотоксическим $CD8^+$ Т-лимфоцитам. Пептиды, связываемые белками MHC класса I, возникают от протеолитической деградации антигенов в цитозоле.

Белки MHC класса II. Одна из преобладающих разновидностей молекул MHC. В большинстве случаев белки MHC класса II осуществляют презентацию пептидов хелперным $CD4^+$ Т-лимфоцитам. Пептиды, связываемые белками MHC класса II, возникают от протеолитической деградации антигенов в эндосомах и лизосомах.

Белки поддержания структуры хромосом (SMC). Группа белков, включающая когезины (удерживающие сестринские хроматиды вместе) и конденсины (участвующие в конденсации хромосом).

Белковый сплайсинг. Аутокаталитический процесс удаления интеина и смыкания экстеинов стандартной пептидной связью.

Белок bHLH. Белок, имеющий основный ДНК-связывающий регион, взаимодействующий со структурой типа «спираль-петля-спираль» (helix-loop-helix).

Белок, связывающий одноцепочечную ДНК. Белок, связывающий одноцепочечную ДНК, предотвращая тем самым образование двуцепочечной структуры.

Бивалент. Структура, объединяющая все четыре хроматиды (по две для каждого из гомологов), формирующаяся в начале мейоза.

Блокированная рамка считывания. Рамка считывания, прерванная стоп-кодоном и неспособная транслироваться в полноразмерный белок.

Бляшка. В лабораторной практике так называют область просветления бактериального газона (сплошной культуры в слое геля); след жизнедеятельности одной фаговой частицы, прошедшей множество циклов литического роста.

Болезнь HbH связана с присутствием в крови большого количества ненормального тетрамера β_4 (относительно небольшого количества нормального гемоглобина $\alpha_2\beta_2$).

Большая бороздка ДНК имеет ширину 22 Å.

Большая субъединица рибосомы (50S у бактерий, 60S у эукариот) имеет активный центр пептидил-трансферазы, который синтезирует пептидную связь.

Быстро останавливающийся мутант. Разновидность мутантов *E. coli*, чувствительных к температуре. При повышении температуры до 42 °С у таких мутантов (*dna*) немедленно останавливается репликация ДНК.

Вариабельная область (V-область). Сайт связывания антигена в молекуле иммуноглобулина или Т-клеточного рецептора. V-области образованы вариабельными доменами полипептидных цепей. Кодируются V-сегментами соответствующих генов, широко варьируют у разных рецепторов (вследствие большого многообразия генов, а также изменений, вносимых при синтезе).

Вегетативная фаза. Период нормального роста и деления бактерии. Для бактерии, способной к споруляции, вегетативная фаза противоположна фазе споруляции, во время которой образуются споры.

Ведущая цепь ДНК. Одна из дочерних цепей, синтезируемая непрерывно в направлении $5' \rightarrow 3'$.

Величина C. Общее количество ДНК в геноме (в расчете на один гаплоидный набор хромосом).

Вирион. Обособленная вирусная частица (безотносительно ее способности заражать клетки и размножаться).

Вироид. Небольшая инфекционная нуклеиновая кислота, которая не имеет белковой оболочки.

Вирулентные мутанты бактериофагов не способны к переходу в лизогенное состояние.

Вирус с дефектом репликации не может совершить инфекционный цикл, поскольку утратил часть необходимых генов.

Вирусное надсемейство транспозонов включает в себя мобильные элементы, родственные ретровирусам. Родство определяется наличием последовательностей, кодирующих обратную транскриптазу или интегразу.

Вирусоид. Небольшая инфекционная нуклеиновая кислота, которая упакована в капсид вируса растений вместе с собственным геномом вируса.

Вирус-помощник. Вирус, который, при смешанной инфекции, обеспечивает дефектному вирусу отсутствующие функции, помогая ему совершить литический цикл.

Внехромосомный геном в бактерии. Обособленная молекула ДНК, ведущая автономное существование внутри бактериальной клетки. Содержит набор генов, необходимый для поддержания собственной репликации; нередко несет гены, необходимые для успешного роста бактерии в определенных условиях среды.

Внеядерные гены находятся за пределами ядра, в органеллах, таких как митохондрии и хлоропласты.

Воблинг-гипотеза (гипотеза «качелей»). Объясняет способность тРНК узнавать больше одного кодона благодаря нестандартному взаимодействию (не G-C и не A-T) с третьим основанием кодона.

Возврат фермента. Восстановление ферментом своей исходной формы для участия в новом акте катализа.

Время удвоения. Период (обычно измеряемый минутами), необходимый для воспроизводства бактериальной клетки.

Врожденный иммунитет. Быстрый иммунный ответ, опосредованный клетками с невариабельными рецепторами (кодируемыми в клетках зародышевого пути), которые распознают патоген. Действие клеток врожденного иммунного ответа направлено на уничтожение патогена и инициацию адаптивного иммунного ответа.

Вставочная последовательность (IS). Небольшой бактериальный транспозон, несущий только гены, которые необходимы ему для транспозиции.

Вторичная реверсия. Случай супрессии эффекта первой мутации второй мутацией.

Вторичный иммунный ответ. Ответ иммунной системы организма при второй встрече с антигеном. По сравнению с первичным ответом характеризуется более быстрой индукцией, большим размахом и более высокой аффинностью антител.

Вторичный сайт прикрепления. Локус бактериальной хромосомы, куда безуспешно пытается встроить себя фаг лямбда (дело в том, что локус содержит сайт, напоминающий сайт *att*).

Выбор копии. Тип рекомбинации, характерный для РНК-содержащих вирусов, при котором РНК-полимераза меняет матрицу во время синтеза.

Вырожденность по третьему основанию. Особенность генетического кода: третий нуклеотид кодона не так существен для его значения, как первые два.

Вышележащие последовательности (по отношению к данному сайту) расположены в направлении, противоположном направлению экспрессии. Например, бактериальный промотор лежит выше транскрипционной единицы, инициирующий кодон лежит выше (предшествует) кодирующей области.

Вышележащие последовательности-активаторы (UAS). Дрожжевой эквивалент энхансера высших эукариот; связывается белками GAL4 (активаторами транскрипции).

Гаплотип. Конкретное сочетание аллелей в определенной области некой хромосомы — можно сказать, «генотип в миниатюре». Первоначально термин использовался для обозначения комбинаций аллелей МНС. В настоящее время он используется более широко, например для обозначения конкретных комбинаций однонуклеотидных замен (ОНЗ), RFLP или других маркеров.

Гаптен. Небольшая молекула, которая действует как антиген, будучи конъюгированной с белком.

Ген. Участок ДНК, задающий полипептидную цепь; помимо собственно кодирующей области, содержит дополнительные последовательности: лидер (предшествует кодирующей области), трэйлер (замыкает кодирующую область), а также интроны (в промежутках между экзонами — разобщенными сегментами кодирующей последовательности).

Генетический код. Соответствие между триплетами нуклеотидов в ДНК (или РНК) и аминокислотами в белке.

Геном. Полный набор последовательностей генетического материала организма. Кроме последовательностей всех хромосом включает в себя последовательности ДНК органелл.

Ген-регулятор кодирует продукт (как правило, белок), контролирующий экспрессию других генов (обычно на уровне тртанскрипции).

Гены «домашнего хозяйства». Гены, (теоретически) экспрессирующиеся в каждой клетке, потому что они обеспечивают жизненно важные функции, необходимые клетке любого типа.

Гены С (гены «си»). Кодируют константные участки полипептидных цепей иммуноглобулина.

Гены роскоши. Гены, кодирующие специализированные белки, синтезируемые, как правило, в больших количествах в конкретных клеточных типах.

Гены теплового шока. Набор локусов, активируемых в ответ на повышение температуры (и других стрессовых факторов). Гены теплового шока есть у всех организмов. Обычно к ним относятся гены шаперонов, работающих с денатурированными белками.

Гетерогенная ядерная РНК (гяРНК) состоит из транскриптов ядерных генов, сделанных РНК-полимеразой II. Характеризуется широким разбросом размеров и низкой стабильностью.

Гетерогенный ядерный РНП (гяРНП). Рибонуклеопротеин на основе гетерогенной ядерной РНК (гяРНК) в комплексе с белками. Пре-мРНК не экспортируются из ядра до тех пор, пока не будет завершен их процессинг; таким образом, гяРНП присутствует только в ядре.

Гетеродуплексная ДНК образуется при взаимодействии комплементарных одиночных цепей, происходящих из разных родительских дуплексов, что имеет место во время генетической рекомбинации.

Гетеромультимер. Белок, состоящий из неидентичных субъединиц (кодируемых разными генами).

Гетерохроматин. Структура, соответствующая высококонденсированному состоянию определенных областей генома, которые не транскрибируются и реплицируются в последнюю очередь. Гетерохроматин бывает двух типов: конститутивный и факультативный.

Гибридизация *in situ*. Метод исследования местоположения нуклеиновых кислот в живой клетке или организме, основанный на добавлении меченой пробы (одноцепочечной ДНК или РНК) к денатурированной ДНК или РНК, находящейся на своем природном месте, в составе препарата клеток или ткани.

Гибридизация. Взаимодействие комплементарных цепей РНК и ДНК с образованием гибридного дуплекса РНК-ДНК.

Гибридный дисгенез. Невозможность скрещивания между определенными линиями *D. melanogaster*: гибриды получаются стерильными (хотя могут выглядеть нормально).

Гибридома. Клеточная линия, образованная от слияния клетки миеломы с лимфоцитом; способна к непрерывной и бесконечной экспрессии иммуноглобулинов обоих «родителей».

Гипермутация. Внесение соматических мутаций в перестроенный ген иммуноглобулина. Эти мутации могут изменять аминокислотную последовательность соответствующего антитела, особенно в сайте связывания антигена.

Гиперчувствительный сайт. Короткий участок хроматина с высочайшей чувствительностью к действию ДНКазы I и других нуклеаз; соответствует области, где (в данный момент) отсутствуют нуклеосомы.

Гипотеза единственной Х-хромосомы. Гипотеза, описывающая инактивацию Х-хромосомы у самок млекопитающих.

Гистон центральной частицы. Одна из четырех разновидностей белков-гистонов (Н2А, Н2В, Н3 и Н4), входящих в состав центральной частицы нуклеосомы. (К гистонам сердцевины не относится гистон Н1.)

Гистон-ацетилтрансферазы (НАТ). Ферменты, которые модифицируют гистоны путем добавления ацетильных групп. Такой активностью обладают некоторые коактиваторы транскрипции.

Гистон-деацетилазы (HDAC). Ферменты, которые удаляют у гистонов ацетильные группы; могут быть ассоциированы с репрессорами транскрипции.

Гистоновая складка. Мотив, присутствующий в каждом из четырех гистонов сердцевины; состоит из трех α-спиралей, соединенных двумя петлями.

Гистоны. Консервативные ДНК-связывающие белки, входящие в состав нуклеосомы — основной субъединицы хроматина эу-

кариот. Гистоны H2A, H2B, H3 и H4 образуют октамер — сердцевину нуклеосомы, на которую намотана ДНК. Гистон H1 располагается снаружи от центральной частицы нуклеосомы.

Главный комплекс гистосовместимости (MHC). Область хромосомы, содержащая гены, задействованные в иммунном ответе. Эти гены кодируют белки для презентации антигена, цитокины, белки комплемента и ряд других функциональных белков. Область MHC отличается высоким полиморфизмом. Гены и белки MHC подразделяются на три класса.

Головная часть. Так называется ДНК-связывающий домен репрессора *lac*.

Гомеодомен. ДНК-связывающий мотив, присущий факторам транскрипции соответствующего класса. Последовательность ДНК, кодирующая этот мотив, известна как гомеобокс.

Гомологичная рекомбинация. Реципрокный обмен последовательностями ДНК, например между двумя хромосомами, несущими одни и те же локусы.

Гомомультимер. Белок, составленный из идентичных субъединиц.

Горячая точка. Геномный сайт, в котором частота мутаций (или рекомбинации) очень сильно повышена как минимум на порядок по сравнению с близлежащими сайтами.

Градиент плотности используется для разделения макромолекул на основании различий в их плотности. Приготавливается из тяжелых растворимых соединений, таких как CsCl.

Группа комплементации. Набор мутаций, которые не дают комплементации в попарных *транс*-сочетаниях. Маркируют генетическую единицу (цистрон).

Группа совместимости плазмид. Совокупность разных плазмид, которые не могут вести совместное существование в одной бактериальной клетке.

Гуморальный ответ. Иммунный ответ, опосредованный главным образом антителами. Иммунитет такого рода может быть передан от одного организма другому инъекцией сыворотки с антителами.

Двугибридный анализ. Метод, позволяющий выявить взаимодействие двух белков за счет определения их способности совмещать в пространстве ДНК-связывающий домен и активирующий (транскрипцию) домен. Анализ выполняют на дрожжах, используя репортерный ген, отвечающий на взаимодействие двух указанных доменов.

Двунаправленная репликация. Репликация, при которой с ориджина в противоположных направлениях стартуют две репликативных вилки.

Двухцепочечный разрыв в ДНК возникает при внесении разрезов в обе цепи дуплекса в одном и том же сайте. Генетическая рекомбинация инициируется при наличии двухцепочечных разрывов. В клетке, не нуждающейся в рекомбинации, действуют системы репарации, направленные на восстановление таких разрывов.

Деацетилаза. Фермент, который удаляет у белков ацетильные группы.

Деацилированная тРНК не несет на себе ни аминокислоты, ни полипептидной цепочки; она выполнила свою миссию и готова покинуть рибосому.

Деградосома представляет собой комплекс бактериальных ферментов, обладающий активностями РНКазы и хеликазы, которые могут быть задействованы в деградации мРНК.

Дезоксирибонуклеаза. Фермент, атакующий связи в ДНК. Может разрезать одну цепь или обе цепи дуплекса.

Дезоксирибонуклеиновая кислота, ДНК. Нуклеиновая кислота, молекулы которой представляют собой длинные полимеры (дезоксирибо)нуклеотидов. В двуцепочечной ДНК две цепи нуклеотидов удерживаются вместе водородными связями между комплементарными азотистыми основаниями.

Делеция. Удаление участка ДНК, сопровождаемое стыковкой фланкирующих последовательностей (за исключением терминальных делеций на концах хромосом).

Деметилаза. Фермент, который удаляет метильные группы. Субстратами различных деметилаз служат ДНК, РНК и белки.

Денатурация белка. Переход белка из конформации, характерной для физиологических условий, в какую-либо другую (неактивную) конформацию.

Дивергенция. Различие, выраженное в процентах, между двумя похожими нуклеотидными последовательностями ДНК или между аминокислотными последовательностями двух родственных белков.

Дицентрическая хромосома. Продукт слияния двух фрагментов хромосомы, каждый из которых имеет центромеру. Продукт нестабилен и может быть разломан, когда две центромеры притягиваются к противоположным полюсам во время митоза.

Длинные концевые повторы (ДКП). Последовательность, присутствующая в одной и той же ориентации на обоих краях провируса (интегрированной последовательности ретровируса).

ДНК, состоящая из повторов. Одна из фракций геномной ДНК. В реакции ренатурации геномных последовательностей ведет себя как скопление множества идентичных (или очень похожих) последовательностей, ренатурирующих между собой по принципу случайного выбора ближайшего соседа.

ДНК-лигаза. Фермент, который образует связь между смежными (3′-OH и 5′-P) концами одноцепочечного разреза (надреза, ника) в дуплексе ДНК.

ДНК-полимераза. Фермент, который синтезирует дочернюю цепь ДНК (по инструкции, исходящей от матричной цепи). Конкретные разновидности этого фермента участвуют в репликации, репарации, или в том и другом.

ДНК-репликаза. ДНК-синтезирующий фермент, необходимый для собственно репликации.

ДНК-фингерпринтинг (генетический «отпечаток пальца»). Методики, которые позволяют выявить различия между особями в распределении длин фрагментов — продуктов расщепления эндонуклеазой рестрикции или продуктов ПЦР. Мишенью реакций служат множественные короткие повторы в геномной ДНК. Точные размеры участков, образованных такими повторами, являются индивидуальными характеристиками; наличие идентичных паттернов говорит о генетическом родстве тестируемых особей (например, отца и сына).

Домен белка. Обособленная непрерывная часть аминокислотной последовательности, которую можно отождествить с конкретной функцией.

Домен хромосомы. Термин употребляется в двух значениях. Он обозначает либо обособленную структурную область хромосомы, где сверхспирализация ДНК происходит независимо, либо обширную область хромосомы, содержащую экспресси-

руемый ген, с повышенной чувствительностью к деградации ферментом ДНКазой I.

Доминантно-негативная мутация. Так говорят, когда продукт мутантного гена препятствует работе продукта гена дикого типа, что приводит к потере или редукции нормальной функции гена в гетерозиготных клетках. Самый распространенный механизм работает для генов, кодирующих гомомультимерные белки, которые неполноценны, если хотя бы одна из субъединиц является продуктом мутантного гена.

Дополнительная шпилька тРНК находится между антикодоновой шпилькой и шпилькой TψC. Эта шпилька тРНК является самой изменчивой по длине, от 3 до 21 основания. Некоторые тРНК вообще не имеют этой шпильки; их относят к классу 1. Остальные тРНК относят к классу 2.

Дочерняя цепь, дочерний дуплекс ДНК. Термины относятся к новосинтезированной ДНК.

ДП-тимоцит, дважды позитивный тимоцит. Незрелая Т-клетка, экспрессирующая поверхностные CD4 и CD8. Селекция ДП-тимоцитов в тимусе приводит к появлению зрелых Т-клеток, экспрессирующих либо CD4, либо CD8.

Единичная клетка. Статус бактериальной клетки (*E. coli*) сразу после деления. Она имеет длину 1,7 мкм и один ориджин репликации.

Задержанно-ранние гены фага лямбда соответствуют средним генам у других фагов. Эти гены не могут быть транскрибированы до того, как будет синтезирован белок-регулятор, кодируемый одним из непосредственно ранних генов.

Замена цепи. Способ репликации некоторых вирусов, при котором рост новой цепи ДНК вытесняет предыдущую (гомологичную) цепь из дуплекса.

Запрограммированный сдвиг открытой рамки считывания необходим некоторым вирусам для экспрессии отдельных белков. Сдвиг +1 или −1 происходит в специальном сайте вирусного генома с определенной заданной частотой.

Затравка (или праймер). Короткая одноцепочечная ДНК или РНК, взаимодействующая с комплементарной одиночной цепью ДНК. ДНК-полимераза использует 3′-конец затравки, чтобы начать свою работу (синтез ДНК на матрице цепи, комплементарной затравке).

Захват экзонов. Способ идентификации сайтов сплайсинга. Произвольный фрагмент геномной ДНК вставляется в специальный вектор. Если фрагмент содержит сайт сплайсинга, вектор начинает вести себя предсказанным образом (например, дает цветную реакцию).

Зоо-блот. Использование саузерн-блоттинга для проверки способности пробы ДНК из одного вида организмов к гибридизации с ДНК из геномов множества других видов.

Избыточность. Случай выполнения одной функции сразу несколькими генами, так что каждый из них в отдельности не существен.

Изошизомеры. Эндонуклеазы рестрикции, узнающие одну и ту же последовательность, чьи эффекты по-разному зависят от статуса метилирования этой последовательности.

Икосаэдрическая симметрия характерна для вирусов, капсиды которых представляют собой многогранники.

Иммунитет у плазмид. Способность плазмиды предотвращать внедрение в клетку другой плазмиды того же типа (обычно за счет вмешательства в ход репликации новой плазмиды).

Иммунитет у транспозонов. Способность некоторых транспозонов предотвращать внедрение в свою молекулу ДНК других транспозонов одного с ними типа.

Иммунитет у фагов. Способность профага предотвращать внедрение в клетку другого фага того же типа (препятствовать инфекции). Является результатом наработки фагового репрессора с генома профага.

Иммунный ответ. Реакция организма на антиген, опосредованная компонентами иммунной системы.

Иммуноглобулин. Белок, продуцируемый В-клетками, который связывает конкретный антиген.

Импринтинг. Термин означает изменение в состоянии гена, происходящее в период образования яйцеклеток или сперматозоидов, в результате которого отцовский и материнский аллели раннего эмбриона обладают разными свойствами. Импринтинг обусловлен дифференциальным метилированием ДНК.

Инвариантные позиции в тРНК содержат одинаковые нуклеотиды почти во всех (>95%) тРНК.

Инверсионная модель. Описывает работу фермента ДНК-гиразы. Фермент связывает положительный сверхвиток (индуцируя образование компенсирующего отрицательного сверхвитка в закрытой кольцевой ДНК). В распоряжении фермента, связавшего сверхвиток, находятся два дуплекса (две дуги сверхвитка). Фермент разрывает обе цепи одного дуплекса, пропускает второй дуплекс через разрыв, склеивает разрыв.

Инвертированные концевые повторы. Короткие идентичные (или очень похожие) последовательности, присутствующие во взаимно-обратной ориентации на краях некоторых транспозонов.

Индуктор. Небольшая молекула, которая запускает транскрипцию генов, связываясь с белком-регулятором.

Индукция. Также означает способность бактерии (или дрожжевой клетки) к синтезу некоторых ферментов лишь при условии наличия их субстратов; применительно к экспрессии генов термин означает включение транскрипции в результате взаимодействия индуктора (которым обычно служит субстрат) с белком-регулятором.

Индукция профага. Вхождение профага в литический (инфекционный) цикл в результате уничтожения лизогенного репрессора, которое приводит к выходу фаговой ДНК из состава бактериальной хромосомы.

Индуцированные мутации возникают под действием мутагена, который может действовать на ДНК прямо (модифицируя ее азотистые основания) либо косвенно (запуская какой-нибудь клеточный путь, ведущий к изменению последовательности ДНК).

Инициация. Начальные этапы транскрипции вплоть до образования первой связи в РНК. К инициации относится связывание РНК-полимеразы с промотором и плавление короткого участка ДНК на одиночные цепи.

Инициирующий кодон. Особый кодон (обычно AUG), с которого начинается синтез белка.

Инсерция, или вставка. Добавление в молекулу ДНК нового участка (или сам добавленный участок). Особой разновидностью вставок являются дупликации.

Инсулятор. Последовательность, препятствующая распространению регуляторного эффекта (активирующего или инактивирующего) вдоль молекулы ДНК.

Интасома. Комплекс, состоящий из белка и ДНК, между интегразой фага лямбда (Int) и сайтом прикрепления фага лямбда (attP).

Интеграза. Фермент, ответственный за сайт-специфичную рекомбинацию, при которой одна молекула ДНК вставляется в другую.

Интеграция, или внедрение. Встраивание вирусной или какой-либо другой последовательности ДНК.в геном хозяина (в качестве вставки, ковалентно пришитой обоими концами к фланкирующим последовательностям хозяина).

Интеин. Часть аминокислотной последовательности, удаляемая у полипептида-предшественника, процессируемого по типу белкового сплайсинга.

Интерференция РНК. Техника манипулирования клеткой на молекулярном уровне. Клетка получает двуцепочечную РНК, идентичную транскриптам конкретного гена-мишени; после этого экспрессия гена-мишени резко затухает. Эффект обусловлен запуском прицельной деградации мРНК гена-мишени.

Интрон. Сегмент гена, который транскрибируется, но впоследствии исключается из транскрипта во время сплайсинга, когда экзоны (сегменты, фланкирующие интрон) смыкаются друг с другом.

Инцизия. Этап работы системы эксцизионной репарации некомплементарностей. На этом этапе эндонуклеаза узнает поврежденный участок ДНК и удаляет его посредством внесения надрезов по обе стороны от повреждения.

Искажение структуры. Изменение конформации ДНК вследствие наличия нуклеотидов или нуклеотидных пар, не вписывающихся в нормальный дуплекс.

Искусственный генетический анализ на чипах. Автоматизированная методика анализа мутантов пекарских дрожжей, при которой анализируемый мутантный штамм скрещивают с упорядоченным массивом из примерно 5000 делеционных мутантов с целью определения возможности появления синтетического летального фенотипа за счет взаимодействия мутаций.

Исправление ошибок синтеза полимера. Какой-либо из механизмов коррекции ошибок синтеза нуклеиновой кислоты или белка, основанных на испытании субъединиц уже после их включения в состав растущего полимера.

Истинная реверсия. Мутация, возвращающая последовательность ДНК к исходному виду.

Истинные терминаторы. Могут терминировать транскрипцию, работу бактериальной РНК-полимеразы, в отсутствие какихлибо дополнительных факторов.

Капсид. Наружная белковая оболочка вирусной частицы.

Каркас хромосомы. Обширная структура, состоящая из белков, с очертаниями пары сестринских хроматид — препарат хромосомы, из которого удалили ДНК и гистоны.

Карта рестрикции. Линейный массив сайтов-мишеней различных эндонуклеаз рестрикции в произвольной ДНК.

Каскад. Последовательность событий, каждое из которых стимулируется предыдущим. Для регуляции транскрипции (как можно наблюдать на примере споруляции или хода литического цикла у фагов) это означает, что регуляция поделена на стадии, на каждой из которых экспрессируется ген, кодирующий регулятор, небходимый для экспрессии генов следующей стадии.

«Кассетная модель» детерминации типа спаривания у дрожжей подразумевает наличие одного активного локуса (активной кассеты) и двух неактивных локусов (молчащих кассет). Замена активной кассеты одного типа молчащей кассетой другого типа приводит к смене типа спаривания дрожжевой клетки.

Катенированные молекулы. Две кольцевые молекулы, соединенные посредством зацепления друг за друга.

«Катящееся кольцо». Тип репликации, при котором репликативная вилка движется вокруг кольцевой матрицы, совершая неопределенно большое число оборотов. По ходу каждого оборота новосинтезированная цепь ДНК смещает с матрицы цепь, синтезированную на предыдущем обороте. Продукт репликации, «хвост», представляет собой линейную серию копий последовательности, комплементарной кольцевой матричной цепи.

кДНК. Одноцепочечная ДНК, комплементарная РНК, синтезированная с нее путем обратной транскрипции in vitro.

Кинетическое исправление ошибок. Механизм исправления ошибок в синтезе полимера. Принцип его действия в том, что скорость совершения «неверных» событий меньше, чем у «верных» событий; поэтому «неверные» события успевают обратиться вспять прежде, чем в полимер будет добавлена новая субъединица.

Кинетохор. Небольшая органелла, ассоциированная с поверхностью центромеры, прикрепляющая хромосому к микротрубочкам веретена деления. Каждая митотическая хромосома имеет два «сестринских» кинетохора, направленных в противоположные стороны.

Кирромицин. Антибиотик; ингибирует синтез белка за счет воздействия на EF-Tu.

Кластер генов. Группа смежных генов, идентичных или родственных друг другу.

Клеверный лист. Сравнение, употребляемое для описания структуры тРНК, изображенной в двух измерениях, в которой четко видны четыре двухцепочечных плеча с петлями на концах.

Клетка Hfr. Бактерия, несущая плазмиду F, встроенную в хромосому. Название Hfr происходит от high frequency recombination (высокая частота рекомбинации) и связано с тем фактом, что хромосомные гены переходят из клеток Hfr в клетки F^- с намного более высокой частотой, чем из клеток F^+.

Клетка памяти. Лимфоцит, который был стимулирован в ходе первичного иммунного ответа на антиген и который быстро активируется при последующем контакте с тем же антигеном. Клетки памяти дают более быстрый ответ на антиген, чем наивные клетки.

Клеточно-опосредованный ответ. Одна из составляющих иммунного ответа; за нее ответственны Т-лимфоциты. Иммунитет такого рода не может быть передан от одного организма другому прививкой антител.

Клональная делеция. Термин означает элиминацию (полное истребление) клональной популяции лимфоцитов (потомков одной клетки). Клональная делеция может быть индуцирована на определенных стадиях дифференцировки лимфоцитов в случае связывания их рецепторов со своим антигеном.

Клональной селекции теория. Основана на предположении, что каждый лимфоцит экспрессирует единственный рецептор

со специфичностью к конкретному антигену и только связывание с этим антигеном стимулирует лимфоцит к пролиферации для участия в устранении этого антигена. Таким образом, антиген «выбирает» лимфоциты, которые будут активированы. Представление о клональной селекции — одно из основных в современной иммунологии.

Клэмп-лоудер. Белковый комплекс из пяти субъединиц, ответственный за наложение на ДНК β-зажима, фиксирующего положение репликативной вилки.

Коактиваторы. Факторы, участники транскрипции, которые не связывают ДНК, зато нужны для взаимодействия ДНК-связывающих активаторов с основным аппаратом транскрипции.

Когезины. Белки, образующие комплекс, который удерживает две сестринские хроматиды в синаптонемном комплексе друг с другом. К когезинам относятся некоторые белки SMC.

Когнатные тРНК узнаются одной и той же разновидностью тРНК-синтетазы; заряжены одной и той же аминокислотой.

Кодирующая область. Часть гена, которая соответствует аминокислотной последовательности полипептида.

Кодирующая цепь ДНК по своей последовательности идентична мРНК, т. е. с применением генетического кода к ее конкретной области, может быть умозрительно транслирована в полипептидную цепь соответствующего белка.

Кодирующий конец. Образуется во время рекомбинации генов иммуноглобулинов и Т-клеточных рецепторов. Кодирующие концы приходятся на терминальные участки разрезанных кодирующих областей V и D(J). От последующего соединения кодирующих концов возникает так называемый «кодирующий стык».

Кодон. Триплет нуклеотидов, транслируемый как конкретная аминокислота или как сигнал терминации.

Коинтеграт. Структура, образованная от слияния двух репликонов, в одном из которых содержался транспозон, а в другом — нет. Коинтеграт содержит две копии транспозона; они находятся на стыках между репликонами и ориентированы как прямые повторы.

Коллинеарное отношение. Отображение 1:1 последовательности нуклеотидных триплетов в последовательность аминокислот.

Компенсация дозы генов. Механизмы, направленные на компенсацию различий между особями разного пола, связанных с разной дозой X-хромосом в их клетках (по одной у самцов и по две — у самок).

Комплекс TIM находится во внутренней мембране митохондрии; он отвечает за транспорт белков из пространства между мембранами внутрь органеллы.

Комплекс TOM находится в наружной мембране митохондрии; он отвечает за транспорт белков из цитозоля в пространство между мембранами.

Комплекс инициации синтеза белка у бактерий состоит из малой рибосомной субъединицы, факторов инициации и инициирующей аминоацил-тРНК, которая взаимодействует с мРНК по инициирующему кодону AUG.

Комплементарность азотистых оснований определяется как способность к взаимодействию *vis-á-vis* в составе двойной спирали, в реакции спаривания (A с T в ДНК, или с U в РНК; C с G).

Комплементация *in vitro*. Метод идентификации компонентов процесса. Экспериментальная система представляет собой реконструкцию нарушенного процесса на основе клеточных экстрактов соответствующего мутанта. Различные фракции клеток дикого типа тестируются на способность к восстановлению активности этой системы.

Комплементация между мутантами. Проявление нормального фенотипа у диплоидной особи, гетерозиготной по каждой из мутаций.

Конверсия гена. Внесение исправлений (замена неспаренных оснований) в одну из цепей гетеродуплексной ДНК для достижения полной комплементарности со второй цепью.

Конденсины. Белки, компоненты комплекса, который связывает хромосомы, чтобы вызвать их конденсацию перед мейозом или митозом. Конденсины относятся к семейству белков SMC.

Консенсусная последовательность. «Идеализированная» последовательность, где на каждой позиции стоит нуклеотид, наиболее характерный для этой позиции (по результатам сравнения между собой большого числа реальных гомологичных последовательностей).

Консервативная транспозиция. Перемещение крупных элементов, первоначально классифицированных как транспозоны, но впоследствии признанных эписомами. По своему механизму такая транспозиция напоминает эксцизию и интеграцию бактериофага.

Консервативные последовательности (нуклеотидные или аминокислотные) содержат одни и те же основания (аминокислоты) в конкретных положениях. Их можно идентифицировать путем сравнения множества последовательностей, объединяемых по какому-то принципу (одинаковые функции, общность происхождения и т.п.)

Константный участок (C-участок) иммуноглобулина или Т-клеточного рецептора представляет собой наименее вариабельную часть белка, примерно одинаковую для всех молекул данного типа. Константные участки кодируются C-сегментами соответствующих генов. Тип иммуноглобулина определяется C-участками тяжелых цепей, которые запускают работу эффекторов того или иного типа.

Конститутивный гетерохроматин. Инертное состояние определенной части геномной ДНК, состоящей из последовательностей, которые никогда не экспрессируются (в основном это — сателлитная ДНК).

Конститутивный процесс происходит все время, не меняясь от воздействия каких-либо стимулов или внешних условий.

Контекст кодона в мРНК. Последовательности, окружающие данный кодон в конкретной мРНК. Термин используется при обсуждении влияния этих последовательностей на эффективность узнавания данного кодона аминоацил-тРНК или на эффективность использования данного кодона как терминатора.

Контролирующие элементы кукурузы. Мобильные элементы, первоначально идентифицированные в качестве генетических единиц. Среди них различают автономные (способные к независимой транспозиции) и неавтономные (способные к транспозиции в присутствии автономного элемента) контролирующие элементы.

Контроль числа кроссоверов. Ограничение рекомбинации между двумя мейотическими хромосомами до одного-двух событий на каждую пару гомологов.

Контр-транскрипт. Молекула РНК, которая мешает РНК-праймеру инициировать транскрипцию, взаимодействуя с ним по принципу комплементарности.

Конъюгация. Процесс, при котором две клетки вступают в контакт для передачи генетической информации. У бактерий ДНК переносится из клетки-донора в клетку-реципиент; у простейших этот поток двусторонний.

Координированная регуляция. Система контроля экспрессии, общая для группы генов.

Копийность. Число копий плазмиды, поддерживаемое в бактериальной клетке (по отношению к числу копий ориджина бактериальной хромосомы).

Кордицепин. 3'-дезоксиаденозин, ингибитор полиаденилирования РНК.

Корепрессор представляет собой небольшую молекулу, которая, связывая белок-регулятор, запускает репрессию трансрипции.

Корончатый галл. Болезнь цветковых растений; опухоль, индуцируемая при заражении бактерией *Agrobacterium tumefaciens*.

Косупрессия. Термин обозначает способность трансгена (как правило, у растений) к ингибированию экспрессии соответствующего эндогенного гена.

Котрансляционная транслокация. Движение белка сквозь мембрану по мере его синтеза. Термин чаще всего употребляется в случаях, когда рибосома связана с каналом; такая форма транслокации характерна для эндоплазматического ретикулума.

Коэффициент сжатия (коэффициент упаковки). Степень компактизации геномной ДНК; рассчитывается как отношение длины молекулы ДНК к длине волокна, в котором она содержится.

Кроссинговер. Реципрокный обмен материалом между гомологичными хромосомами, который происходит в профазе I мейоза и обеспечивает генетическую рекомбинацию.

Куру. Неврологическое заболевание человека, вызываемое прионами. Распространена у каннибалов; передается при поедании инфицированного мозга.

Кэп 0 метилирован лишь по положению 7 в гуанине.

Кэп 1 метилирован по двум положениям (7 в гуанине кэпа и 2'-O следующего нуклеотида).

Кэп 2 метилирован по тем же положениям, что и кэп 1 плюс еще по положению N6 в аденине на 5'-конце мРНК.

Кэп. Представляет собой структуру на 5'-конце эукариотической мРНК. Вносится после транскрипции, посредством сшивки концевого фосфата 5'-ГТФ с концевым основанием мРНК. Добавленный гуанин (иногда — другое основание) метилирован, так что вся структура имеет форму 7MeG5'ppp5'Np.

Латеральный элемент. Структура в составе синаптонемного комплекса, которая образуется по мере конденсации сестринских хроматид вдоль аксиального элемента.

Легкая цепь иммуноглобулина. Один из двух типов субъединиц тетрамера, которым является антитело. Каждое антитело содержит две легких цепи. N-конец легкой цепи образует часть сайта узнавания антитела.

Лейциновая молния. Мотив, обеспечивающий димеризацию; характерен для представителей целого класса факторов транскрипции.

Лидерная 5'-концевая нетранслируемая область мРНК. Нетранслируемая 5'-концевая последовательность, предшествующая инициирующему кодону.

Лидерная последовательность белка. Короткая N-концевая последовательность, обеспечивающая инициацию прохождения полипептида в мембрану, или сквозь нее.

Лидерный пептид. Продукт трансляции короткой кодирующей последовательности, используемый для регуляции транскрипции триптофанового оперона через контроль движения рибосомы.

Лизис. Смерть бактерии, завершающая инфекционный цикл бактериофага. Во время лизиса бактерия лопается (потому что фаг обеспечивает ферменты, которые разъедают плазматическую мембрану или клеточную стенку). Погибшая клетка высвобождает множество зрелых инфекционных частиц фага. Термин также используется в отношении эукариотических клеток (например, в отношении инфицированных клеток, атакуемых иммунной системой, которые подвергаются лизису).

Лизогения. Способность фага населять бактериальную клетку в качестве профага — стабильного компонента бактериального генома.

Липополисахарид (LPS). Молекула, содержащая липидный и углеводный компоненты. Присутствует на наружной мембране грам-отрицательных бактерий. Работает как эндотоксин, индуцирующий септический шок во время инфекции.

Литическая инфекция бактерии бактериофагом заканчивается гибелью бактерии с высвобождением из нее множества потомков бактериофага.

Лицензирующий фактор. Находится в ядре, необходим для репликации; инактивируется или разрушается после одного раунда репликации. Для последующих раундов репликации необходим приток новой порции лицензирующего фактора.

Лишенный TATA-бокса промотор. Промотор без TATA-бокса выше точки начала транскрипции.

Локус H2 в кластере главного комплекса гистосовместимости (MHC) мыши. Кластер генов на хромосоме 17, чьи гены кодируют белки для презентации антигенов, цитокины и белки комплемента.

Локус HLA в главном комплексе гистосовместимости (MHC) человека. Кластер генов на хромосоме 6. Гены *HLA* кодируют белки для презентации антигенов, цитокины и белки комплемента.

Локус. Положение на хромосоме, в котором находится ген, ответственный за конкретный признак. В локусе может находиться любой из аллелей данного гена.

Малая бороздка ДНК имеет ширину 12 Å.

Малая субъединица рибосомы. Связывающая мРНК субъединица рибосомы с коэффициентом седиментации 30S (для бактерий) или 40S (для эукариот).

Малые РНК. Небольшие бактериальные РНК, участвующие в регуляции экспрессии генов.

Малые цитоплазматические РНК (мцРНК). Присутствующие в цитоплазме (и иногда обнаруживаемые в ядре) малые РНК.

Малые ядерные РНК (мяРНК). Один из нескольких видов малых РНК, обнаруживаемых только в ядре. Некоторые из мяРНК участвуют в процессах сплайсинга и ряде других реакций процессинга РНК.

Материнское наследование. Преобладание у потомства генетических маркеров, полученных от конкретного родителя (матери).

Матричная РНК (мРНК). Служит посредником в реализации генетической информации. В ней отображена одна из цепей гена, кодирующего белок. Кодирующая область мРНК транслируется в аминокислотную последовательность белка в соответствии с триплетным генетическим кодом.

Матураза. Белок, кодируемый интроном группы I или II, который помогает РНК принять активную конформацию, необходимую для аутокаталитического сплайсинга.

Медиатор. Большой белковый комплекс, который у бактерий и дрожжей ассоциирован с РНК-полимеразой II. Содержит факторы, необходимые для транскрипции с многих (если не большинства) промоторов.

Межаллельная комплементация. Изменение в свойствах гетеромультимерного белка в результате взаимодействия его субъединиц — продуктов двух мутантных аллелей. Такой белок может быть более или менее активен по сравнению с белком, состоящим из субъединиц одного (первого или второго) типа. Другой причиной межаллельной комплементации может быть трансвекция.

Межцистронный участок. Расстояние между терминирующим кодоном одного гена и инициирующим кодоном следующего гена.

Метилаза *de novo*. Фермент, который добавляет метильную группу к неметилированной последовательности-мишени в ДНК.

Метилтрансфераза. Фермент, который добавляет метильную группу к субстрату (небольшой молекуле, белку или нуклеиновой кислоте).

Миграция цепи. Процесс, в ходе которого одиночная цепь ДНК, частично спаренная с комплементарной цепью другого дуплекса, увеличивает регион спаривания за счет вытеснения из дуплекса исходной цепи.

Микрококковая нуклеаза. Эндонуклеаза, которая расщепляет ДНК; в хроматине отдает предпочтение участкам между нуклеосомами.

МикроРНК. Очень короткие РНК, которые, по-видимому, регулируют экспрессию генов.

Микросателлитная ДНК состоит из очень коротких (<10 пн) повторов.

Мини-клетка. Безъядерная клетка *E. coli*, не имеет хромосомы, продукт аномального деления.

Минимальный (кор) фермент. Комплекс субъединиц РНК-полимеразы, необходимый для элонгации. В комплекс не входят дополнительные субъединицы или факторы, которые могут быть нужны для инициации или терминации.

Мини-сателлитная ДНК состоит из ~10 копий короткой последовательности (длина повтора измеряется десятками пн). Число повторов в разных геномах отличается.

Минус-цепь ДНК. Одноцепочечная ДНК; комплементарна геному ретровируса, представляющему собой одиночную плюс-цепь РНК ((+)-цепь РНК).

Миссенс-мутации. Приводят к альтернативному прочтению одного из кодонов открытой рамки считывания; в результате происходит замена одной аминокислоты на другую.

Миссенс-супрессор представляет собой ген тРНК с мутацией, переключающей тРНК на узнавание миссенс-кодона. Транслируя миссенс-кодон прежней аминокислотой, мутантная тРНК супрессирует эффект миссенс-мутации.

Множественные аллели. Для некоторых локусов известно больше двух аллелей; каждый аллель может характеризоваться собственным фенотипом.

Модификация ДНК или РНК. Любые изменения, сообщаемые нуклеотидам после их включения в полимер.

Модифицированные основания. Любые азотистые основания в составе ДНК, за исключением стандартных (T, C, A, G в ДНК; U, C, A, G в РНК). Модифицированные основания появляются в результате постсинтетических реакций.

Мозаицизм фенотипа. Следствие появления изменений в генотипе во время соматического индивидуального развития.

Мозаичность, обусловленная эффектом положения. Неоднородный фенотипический эффект подавления экспрессии гена под влиянием близости гетерохроматина.

Молекулярный паттерн, ассоциированный с патогеном (PAMP). Молекулярная структура, воспринимаемая иммунной системой как индикатор поверхности патогена. Один паттерн может быть характерен для большого числа патогенов. При иммунном ответе такие паттерны могут распознаваться рецепторами на поверхности клеток врожденного иммунитета.

Молчащие мутации. Мутации, не меняющие последовательность аминокислот в белке, поскольку создают синонимичный кодон.

Молчащий сайт. Измененная позиция в кодирующем регионе гена, не затронувшая последовательность аминокислот в белке.

Моноцистронная мРНК кодирует один белок.

Мпн (мега-пара нуклеотидов). Используется для измерения длины ДНК, равна 10^6 пар нуклеотидов.

Мультикопийный контроль репликации плазмиды позволяет этой плазмиде быть представленной множеством копий на одну бактериальную клетку.

Мутагены повышают частоту мутаций, прямо или косвенно индуцируя изменения в последовательности ДНК.

Мутант с замедленной остановкой репликации. Вариант чувствительного к температуре мутанта *E. coli* по системе репликации, способный завершить уже начатый раунд репликации при непермиссивной температуре, однако неспособный начать новый раунд репликации.

Мутанты *dna* у бактерий чувствительны к температуре: они способны к синтезу ДНК при 37°C, но не способны к синтезу ДНК при 42°C.

Мутатор. Мутация или мутантный ген, повышающий частоту мутаций. Гены-мутаторы нередко кодируют белки, задействованные в репарации поврежденной ДНК.

Мутация с обретением функции. При таких мутациях происходит увеличение нормальной активности гена. Иногда это сопряжено с фенотипическим проявлением каких-то ненормальных свойств. Большинство таких мутаций доминантны.

Мутация с потерей функции. Мутация, из-за которой активность гена ослаблена или сведена на нет. Большинство таких мутаций рецессивны.

Мутация с сохранением остаточного уровня экспрессии оставляет гену часть его функции. Бывает, что мутантный белок частично активен (это характерно для миссенс-мутаций) или что случайные редкие проходы рибосомы через нонсенс-мутацию позволяют получить небольшое количество нормального белка.

Наводящая РНК. Небольшая молекула РНК, комплементарная редактируемой молекуле. Используется в качестве матрицы для изменения последовательности (вставок и делеций нуклеотидов) редактируемой РНК.

Надсемейство генов. Набор гомологичных генов, существенно отличающихся по последовательности.

Наполовину метилированный сайт. Симметричная последовательность (палиндром), метилированная лишь по одной из цепей ДНК.

Наружный домен. Часть белка, закрепленного в плазматической мембране, выходящая в наружную среду клетки. После интернализации такого белка его наружный домен оказывается обращенным в просвет органеллы (топологический эквивалент наружной среды клетки).

Насцентный белок еще не полностью синтезирован; полипептидная цепь все еще связана с рибосомой через тРНК.

Неавтономный контролирующий элемент. Транспозон кукурузы; кодирует нефункциональную транспозазу; может перемещаться по геному лишь в присутствии *транс*-действующего автономного элемента из того же семейства.

Неаллельные гены. Копии одного гена, находящиеся в разных местах генома, — в отличие от аллелей, которые являются копиями одного и того же гена, получены от разных родителей и занимают одни и те же положения на гомологичных хромосомах.

Невирусное надсемейство транспозонов возникло независимо от ретровирусов.

Негативная комплементация. Случай межаллельной комплементации, когда в мультимерном белке мутантная субъединица супрессирует активность субъединицы дикого типа.

Негативная регуляция. Тип регуляции экспрессии генов, которые активны по умолчанию. Выключение такого гена требует особого вмешательства.

Негистоновый белок. Любой структурный белок хромосомы, за исключением гистонов.

Негомологичное соединение концов. Представляет собой лигирование тупых концов ДНК. Характерно для многих путей репарации и для некоторых случаев рекомбинации (например, в генах иммуноглобулинов).

Недокрученная ДНК (ДНК с отрицательной сверхспирализацией). ДНК, имеющая меньше пар нуклеотидов на виток, чем обычные 10 пн. Это означает, что две цепи ДНК меньшее число раз зацеплены друг за друга, в пределе приводя к расхождению цепей ДНК (денатурации).

Независимые индукторы. Аналоги аутентичных индукторов транскрипции, не являющиеся субстратами для индуцированных ферментов.

Неиндуцируемый мутант. Мутант, для которого поврежденный ген (или гены) не может быть экспрессирован (ни при каких условиях).

Нейтральная мутация не дает эволюционных преимуществ и, как правило, не имеет фенотипического эффекта.

Нейтральные замены в белке. Замены отдельных аминокислот, не влияющие на активность белка.

Некомплементарность. Сайт ДНК, где противолежащие азотистые основания двух цепей дуплекса не образуют стандартную пару G-C или A-T. Наличие некомплементарности может быть вызвано включением неверного нуклеотида при репликации либо мутацией одного из оснований.

Непосредственно-ранние гены фага лямбда. Эквивалент ранних генов других бактериофагов. При заражении эти гены немедленно транскрибируются РНК-полимеразой хозяина.

Непродуктивная перестройка. Частный случай рекомбинации V-, (D)-, J-генных сегментов, не приводящий к образованию правильной открытой рамки считывания. Непродуктивная перестройка — результат добавления или вычитания лишнего нуклеотида, прерывающего рамку (хотя некоторые непродуктивные перестройки дают рамку нормальной длины, кодирующую нефункциональный белок).

Непрямое мечение концов. Метод исследования организации ДНК: в конкретный сайт вносится разрез, после чего проводится идентификация всех фрагментов, содержащих последовательность, находящуюся на одном из краев разреза. Позволяет определить расстояние от разреза до следующих разрывов в ДНК.

Нерепликативная транспозиция. Перемещение транспозона, который покидает сайт-донор (обычно оставляя в нем двуцепочечный разрыв), чтобы занять новый сайт.

Нереципрокная рекомбинация. Случается при неверно установленном соответствии и кроссинговере между гомологичными последовательностями ДНК. В результате образуются одна рекомбинантная молекула с делецией и одна рекомбинантная молекула с дупликацией.

Неточная эксцизия. Случай, когда транспозон вырезает самого себя из ДНК, оставляя в ней часть своей последовательности.

Нетранскрибируемый спейсер. Участок между единицами транскрипции в тандемном кластере генов.

Нижележащее местоположение объекта в ДНК по отношению к другому объекту в той же молекуле. Термин задает направление экспрессии: например кодирующая область является нижележащей по отношению к инициирующему кодону и т.п.

Ник-трансляция. Реакция, катализируемая ДНК-полимеразой I *E. coli*. Фермент способен использовать одноцепочечный разрыв в ДНК в качестве стартовой точки деградации одной из цепей дуплекса и ее замещения новым синтезируемым материалом. Применяется, в частности, для внесения метки в ДНК *in vitro*.

Нонсенс-мутация. Изменение последовательности ДНК, в результате которого один из значимых кодонов открытой рамки считывания превращается в терминирующий кодон (UAG, UGA или UAA).

Нонсенс-супрессор. Ген тРНК, содержащий мутацию, приводящую к тому, что тРНК может интерпретировать терминирующий кодон как аминокислоту.

Нопалиновые плазмиды. Разновидность плазмиды Ti, живущей в *Agrobacterium tumefaciens*. Содержит гены для синтеза одного из опинов, нопалина. Возбуждает у растений опухоли, дифференцированные наподобие ранних эмбриональных структур (разновидность корончатого галла).

Нуклеиновые кислоты. Молекулы, носители генетической информации. Представляют собой последовательности азотистых оснований на молекулах рибозы, соединенных фосфодиэфирными связями. ДНК — это дезоксирибонуклеиновая кислота, РНК — рибонуклеиновая кислота.

Нуклеоид. В прокариотической клетке — структура, содержащая геном. Состоит из ДНК в комплексе с белками; аналог эукариотической хромосомы. Нуклеоид не отделен от цитоплазмы мембранной оболочкой.

Нуклеосома. Основная структурная единица хроматина; состоит из ~200 пн ДНК и октамера гистонов.

Нуль-мутация полностью элиминирует функцию гена.

Обмен цепями. Процесс, в ходе которого одна из цепей ДНК-дуплекса уходит и образует дуплекс с комплементарной цепью другой молекулы ДНК, вытесняя при этом исходную цепь.

Обратная мутация. Мутация, обращающая эффект исходной мутации, инактивировавшей ген (и таким образом возвращающая дикий фенотип).

Обратная транкриптаза. Фермент, который использует одноцепочечную РНК в качестве матрицы для синтеза двухцепочечной ДНК.

Обратная транскрипция. Синтез ДНК на матрице РНК. Выполняется ферментом обратной транскриптазой.

Однокопийная репликация. Система контроля репликации, обеспечивающая наличие единственной копии репликона на бактериальную клетку. Присуща бактериальной хромосоме и некоторым плазмидам.

Однонаправленная репликация. Движение репликативной вилки только в одном направлении (от ориджина репликации).

Однонуклеотидный полиморфизм. Межиндивидуальные различия последовательности (полиморфизм), состоящие в изменении единственного нуклеотида. Составляют большую часть генетических вариаций между отдельными индивидами внутри вида.

«Око репликации». Участок реплицированной ДНК внутри более протяженного нереплицированного участка.

Октопиновые плазмиды *Agrobacterium tumefaciens* содержат гены, обеспечивающие синтез опинов октопинового типа. Вызывают образование недифференцированных опухолей (разновидность корончатого галла цветковых растений).

Опал-кодон. Триплет UGA, один из трех кодонов, терминирующих синтез белка. У небольшого числа низших организмов этот кодон соответствует аминокислоте.

Оператор. Сайт ДНК, с которым связывается белок-репрессор, чтобы мешать инициации транскрипции с близлежащего промотора.

Оперон. Единица экспрессии и регуляции бактериальных генов; помимо структурных генов, включает в себя контрольные элементы, узнаваемые продуктами генов регуляторов.

Опин. Производная аргинина; синтезируется в клетках растений, зараженных корончатым галлом.

Опосредованное нонсенс-мутацией разрушение мРНК. Путь деградации мРНК в случае, если стоп-кодон встречается до начала последнего экзона.

Определение интрона. Процесс узнавания пары сайтов сплайсинга в ходе взаимодействий при участии только 5′-сайта и точки ветвления/3′-сайта.

Определение экзона. Процесс опознания пары сайтов сплайсинга при помощи взаимодействий с участием 5′-концевых областей двух соседних интронов.

Ориджин. Последовательность ДНК, на которой инициируется репликация.

Ортологи. Белки из разных организмов, между которыми установлено соответствие на основании гомологии аминокислотных последовательностей.

Ослабляющая мутация в промоторе понижает эффективность транскрипции.

Основной аппарат транскрипции. Комплекс транскрипционных факторов, собирающийся на промоторе до посадки на него РНК-полимеразы.

Основной уровень. Уровень ответа системы в отсутствие каких-либо стимулов. (Основной уровень транскрипции некоторого гена — это уровень, обнаруживаемый в отсутствие какой-либо специфичной активации).

Остановка рибосомы. Ингибирование движения рибосомы, достигшей кодона, для которого отсутствует соответствующая заряженная аминоацил-тРНК.

Отжиг ДНК. Образование двухцепочечной структуры из комплементарных одиночных цепей ДНК.

Открытая рамка считывания. Последовательность нуклеотидов, разбитая на триплеты для трансляции в последовательность аминокислот. Начинается инициирующим кодоном, заканчивается терминирующим кодоном.

Открытый комплекс. Стадия инициации транскрипции, когда РНК-полимераза заставляет цепи ДНК разделиться и образовать «транскрипционный пузырек».

Отстающая цепь. Дочерняя цепь ДНК, которая, согласно направлению движения репликативной вилки, должна расти в неестественном направлении 3′→5′. Отстающая цепь синтезируется в форме небольших фрагментов (5′→3′), впоследствии соединяемых ковалентными сшивками.

Охра-кодон. Триплет UAA, один из трех кодонов, терминирующих синтез белка.

Ошибочно спаренные основания. Некомплементарное взаимодействие двух оснований (например, спаривание аденина с цитозином или тимина с гуанином).

Парадокс величины С заключается в отсутствии взаимосвязи между содержанием ДНК и ее кодирующим потенциалом.

Паранемичное соединение. Участок двуцепочечной молекулы ДНК, на котором две комплементарные цепи не образуют двойной спирали, а просто располагаются бок о бок.

Пептидил-трансфераза. Фермент, который синтезирует пептидную связь при добавлении аминокислоты в растущую полипептидную цепочку. Этот фермент — не что иное, как большая субъединица рибосомы. Собственно каталитическая активность принадлежит рРНК.

Пептидил-тРНК. Молекула тРНК, которая несет на себе растущий полипептид, перенесенный на нее при предыдущем синтезе пептидной связи, во время синтеза белка.

Первичный иммунный ответ. Ответ организма при первой встрече с антигеном. По сравнению с вторичным иммунным ответом, характеризуется меньшей продолжительностью и относительно низкой аффинностью антител.

Первичный транскрипт. Исходный немодифицированный продукт транскрипции; РНК, соответствующая транскрипционной единице.

Переключение класса иммуноглобулина. Изменение в организации гена *Ig*, при котором изменяется С-участок тяжелой цепи, а V-участок остается прежним.

Перекодировка. Присвоение отдельным кодонам или сериям кодонов альтернативных значений вопреки генетическому коду. Перекодировка может быть обусловлена нестандартным взаимодействием аминоацил-тРНК с мРНК под влиянием со стороны рибосомы.

Перекрученная ДНК имеет число нуклеотидных пар на виток, большее обычных десяти. Это означает, что в такой ДНК цепи

скручены более туго друг относительно друга, создавая напряжение внутри молекулы.

Переход одиночной цепи. Реакция, катализируемая топоизомеразой I, в ходе которой сегмент одной из цепей ДНК проходит сквозь разрыв во второй цепи.

Периплазма (или периплазматическое пространство). Пространство между наружной и внутренней мембранами клеточной оболочки грамотрицательных бактерий.

Перисептальное кольцо. Опоясывающая клетку область клеточной оболочки, где наружная и внутренняя мембраны выглядят слившимися. Задает положение септы.

Пероксины. Белковые компоненты пероксисомы.

Пероксúсома. Цитоплазматическая органелла, окруженная одной мембраной. Содержит окислительные ферменты.

Петля. Одноцепочечный участок, которым замыкается шпилька в РНК (или в одноцепочечной ДНК); соответствует участку между инвертированными повторами в двуцепочечной ДНК.

Пили (или ворсинки). Выросты на поверхности бактерии, позволяющие ей прикрепляться к другим бактериям. Похожи на короткие, тонкие, гибкие палочки. В частности, ворсинки используются при конъюгации для переноса ДНК между бактериальными клетками.

Пилин. Белковая субъединица, полимер которой образует ворсинки у бактерий.

Пиримидиновый димер образуется в результате образования прямой ковалентной связи между соседними пиримидинами ДНК. Пиримидиновые димеры образуются при ультрафиолетовом облучении, они блокируют репликацию и транскрипцию ДНК.

Плазмида F. Эписома, которая может существовать отдельно от бактериальной хромосомы либо быть интегрированной в нее; любая из этих форм поддерживает конъюгацию.

Плазмида Ti. Эписома бактерии *Agrobacterium tumefaciens*; несет гены, обеспечивающие индукцию корончатого галла в инфицированных растениях.

Плазмида. Кольцевая внехромосомная ДНК. Автономно существует в цитоплазме бактерии; контролирует собственную репликацию.

Плазмиды Ri обнаружены в *Agrobacterium tumefaciens.* Они похожи на плазмиды Ti и тоже несут гены, которые вызывают у растений развитие заболевания, напоминающего корончатый галл.

Плектонемичное соединение. Участок двуцепочечной молекулы ДНК, на котором ее цепи образуют двойную спираль (в отличие от паранемичного соединения, где двойная спираль отсутствует).

Плечо тРНК. Одна из четырех (в отдельных случаях — пяти) структур типа «стебель-петля», формирующих вторичную структуру тРНК.

Плечи сайта присоединения фага лямбда. Последовательности, фланкирующие центральный регион (регион прохождения рекомбинации).

Плотное связывание РНК-полимеразы с ДНК. Образование открытого комплекса (в котором цепи ДНК разделены).

Плюс-цепочечный вирус распространяет свой геном в форме одноцепочечной нуклеиновой кислоты, которая является смысловой, то есть непосредственно кодирует белки.

Плюс-цепь ДНК. Одна из цепей ДНК ретровируса, идентичная геномной РНК ретровируса (другая цепь дуплекса, минус-цепь, комплементарна геномной РНК ретровируса).

Подверженный ошибкам синтез ДНК. Синтез, двуцепочечный продукт которого содержит множество некомплементарностей из-за включения множества неверных нуклеотидов в дочернюю цепь.

Поддерживающая метилаза добавляет метильную группу к сайту-мишени, который уже наполовину метилирован.

Поздние гены бактериофага транскрибируются после начала репликации фаговой ДНК. Они кодируют компоненты инфекционных частиц.

Поздняя инфекция. Часть литического цикла бактериофага, от начала репликации до лизиса клетки. В ходе поздней инфекции, параллельно репликации фаговой ДНК, идет наработка структурных компонентов инфекционных частиц.

Позиционирование нуклеосом. Упорядоченное положение нуклеосом относительно определенных последовательностей ДНК.

Позиционная информация. Дифференциальное распределение макромолекул в яйцеклетке и раннем эмбрионе. Локализация макромолекул в яйцеклетке сама по себе является одной из форм наследуемой информации.

Поли(A)$^-$ мРНК. Фракция мРНК, не имеющей поли(A).

Поли(A)$^+$ мРНК. Фракция мРНК, имеющей поли(A).

Поли(A). Полимер адениловой кислоты длиной ~200 оснований, добавляемый к 3′-концу эукариотической мРНК после ее синтеза.

Поли(A)-полимераза (PAP). Фермент, который добавляет поли(A) к 3′-концу эукариотической мРНК. Не использует матрицу.

Поли(A)-связывающий белок. Белок, который связывает поли(A) на 3′-конце эукариотической мРНК.

Полиморфизм (генетический полиморфизм). Одновременное присутствие в популяции аллелей с различиями в нуклеотидной последовательности. Сначала учитывались только аллели, дающие разные фенотипы; в настоящее время в качестве полиморфизмов учитываются любые изменения в последовательности гена, даже минимальные нейтральные. В практических целях, изменение нуклеотидной последовательности гена считается полиморфизмом, если это изменение встречается в популяции с частотой >1%.

Полиморфизмы длин фрагментов рестрикции. Наследственные различия в паттерне сайтов рестрикции (могут быть связаны с однонуклеотидными заменами в некоторых сайтах). Используются для отождествления геномных последовательностей с генетическими маркерами.

Полирибосома. мРНК, транслируемая одновременно несколькими рибосомами.

Политенные хромосомы образуются в результате серии репликаций набора хромосом без разделения реплик.

Полицистронная мРНК содержит кодирующие области двух или более генов.

Полностью метилированный сайт. Симметричная последовательность (палиндром), метилированная по обеим цепям ДНК.

Полосы на политенных хромосомах. Регионы с высокой плотностью, содержащие преимущественно ДНК, несущую активные гены.

Полуконсервативная позиция в нуклеотидной последовательности занята всегда одним и тем же типом основания (пурином или пиримидином); вывод делается на основании сравнения множества гомологичных последовательностей.

Полуконсервативная репликация. Выполняется с разделением цепей исходного (родительского) дуплекса, каждая из которых становится матрицей для синтеза комплементарной цепи.

Полуметилированная ДНК. Метилирована по одной из цепей последовательности-мишени (при этом вторая цепь тоже содержит цитозин).

Полунепрерывная репликация. Способ репликации, при котором одна дочерняя цепь синтезируется как непрерывная прямая, а другая — в виде коротких фрагментов.

Полярность. Свойство некоторых мутаций влиять на экспрессию нижележащих генов в пределах транскрипционной единицы (на уровне транскрипции или трансляции).

Последовательность –10. Консенсусная последовательность, расположенная на расстоянии около 10 пн выше точки начала транскрипции бактериального гена. Данная последовательность вовлечена в процесс денатурации ДНК при инициации транскрипции.

Последовательность –35. Консенсусная последовательность, расположенная на расстоянии около 35 пн выше точки начала транскрипции бактериального гена. Вовлечена в процесс первичного узнавания промотора РНК-полимеразой.

Последовательность Шайна–Дальгарно. Полипуриновая последовательность AGGAGG, предшествующая инициирующему кодону AUG в бактериальной мРНК. Комплементарна 3′-концевой последовательности 16S рРНК.

Постмейотическая сегрегация. Разделение двух цепей рекомбинантного дуплекса, не полностью совпадающих по нуклеотидной последовательности (как следствие образования гетеродуплекса при гомологичной рекомбинации) в ходе последующей репликации.

Пострепликационный комплекс состоит из белков ORC, связывающих ДНК ориджина; обнаружен у дрожжей *S. cerevisiae*.

Посттрансляционная транслокация. Движение белка через мембрану после завершения его синтеза и высвобождения из рибосомы.

Правило n–1. Правило утверждает, что в клетках самок млекопитающих активна только одна X-хромосома; все остальные инактивируются.

Правило GT-AG. Означает наличие именно этих динуклеотидов на первых двух и последних двух позициях интронов в ядерных генах.

Праймаза. Разновидность РНК-полимеразы, синтезирующая короткие фрагменты РНК, предназначенные для использования в качестве затравок при репликации ДНК.

Праймосома. Белковый комплекс, инициирующий репликацию в ориджинах типа φX. Также используется для разблокирования застрявших репликативных вилок.

Пре-белок. Так называют белок, предназначенный для импорта в органеллу или для секреции во внешнюю среду бактерии — до тех пор, пока у него не будет удалена сигнальная последовательность.

Предрепликационный комплекс. Комплекс, состоящий из ДНК и белков; находится в ориджине *S. cerevisiae*; необходим для репликации. В его состав входят комплекс ORC, Cdc6, белки MCM.

Представленность мРНК. Среднее число молекул мРНК на клетку.

Представленные мРНК. Несколько видов транскриптов, представленных в клетке в огромном количестве.

Предынициаторный комплекс. В эукариотической транскрипции: сборка факторов транскрипции на промоторе до прихода РНК-полимеразы.

Преждевременная терминация. Прекращение синтеза белка или РНК раньше естественного завершения; в результате получается укороченный абортивный продукт. Преждевременная терминация синтеза белка может быть связана с появлением терминирующих кодонов в открытой рамке считывания (в результате мутации). Преждевременную терминацию синтеза РНК связывают с различными воздействиями на РНК-полимеразу.

Пре-мРНК. Ядерный транскрипт, подлежащий процессингу (модификации и сплайсингу); будущая мРНК.

Приобретенный иммунитет. Ответ организма на антиген, опосредованный специфичной активацией лимфоцитов за счет их взаимодействия с антигеном. Иммунный ответ развивается через несколько дней после того, как лимфоциты с антиген-специфичным рецептором оказываются простимулированы к пролиферации и превращению в эффекторные клетки. Данный процесс отвечает за формирование иммунологической памяти.

Прион. Инфекционный агент белковой природы. Ведет себя как наследуемый признак, несмотря на то что никак не связан с нуклеиновыми кислотами. Примерами прионов служат PrP^{Sc} (агент, вызывающий скрэпи овец и бычий губчатый энцефалит) и Psi (ответственный за особое наследуемое состояние дрожжевых клеток).

Провирус. Двухцепочечная ДНК-копия РНК-генома ретровируса, встроенная в эукариотический геном (фактически область эукариотического генома).

Продуктивная перестройка. Успешное завершение рекомбинации V-, (D)-, J-генных сегментов с попаданием всех сегментов в одну открытую рамку считывания.

Прозрачная бляшка. Содержит только остатки бактериальных клеток, подвергшихся лизису.

Промежутки между полосами. Бледно окрашенные участки поперечно исчерченных политенных хромосом, перемежаемые интенсивно окрашенными полосами.

Промотор. Участок ДНК, с которым связывается РНК-полимераза для инициации транскрипции.

Протеом. Полный набор белков — всех возможных продуктов экспрессии генов целого генома. Размер протеома больше, чем количество генов, потому что некоторые гены кодируют множества белков. Иногда термин «протеом» используется в несколько ином смысле и означает набор белков, экспрессируемых одной клеткой в конкретный момент времени.

Профаг. Геном бактериофага, встроенный в геном бактерии; представляет собой линейную область бактериальной хромосомы.

Процессивность. Способность фермента совершать длинную серию каталитических циклов на одной матрице (вместо того, чтобы диссоциировать после каждого цикла).

Процессинг РНК. Изменения, происходящие с РНК после транскрипции (модификация 5′- и 3′-концов, метилирование внутренних сайтов, сплайсинг).

Процессированный псевдоген. Неактивная копия гена, не содержащая интронов (в отличие от своего прообраза — активного гена с прерывистой структурой). Появление таких псевдогенов объясняется обратной транскрипцией мРНК и последующей вставкой ДНК-копии в геном.

Прямые мутации. Инактивируют ген дикого типа.

Прямые повторы. Идентичные (или очень похожие) последовательности, которые присутствуют в двух или более копиях, в одной и той же ориентации внутри одной молекулы ДНК.

Псевдогены. Неактивные, но стабильные компоненты генома. Псевдоген — это бывший активный ген, в котором однажды произошла мутация, блокировавшая его транскрипцию (или трансляцию, или то и другое).

Пуромицин. Антибиотик, который преждевременно терминирует синтез белка. Благодаря сходству с тРНК, молекула пуромицина успешно ее имитирует, принимая на себя растущую полипептидную цепь.

Пуф. Вздутие полосы политенной хромосомы, связанное с синтезом РНК в каком-то локусе этой полосы.

Разделение коинтеграта. Происходит в результате гомологичной рекомбинации между копиями транспозона в коинтеграте. Продуктами реакции являются репликоны донора и мишени, свободные друг от друга, причем каждый несет копию транспозона.

Разрыв и воссоединение. Вариант рекомбинации, при котором две молекулы ДНК обмениваются участками (идет с формированием протяженного гетеродуплексного региона в районе сайта рекомбинации).

Разрывно-гибридный цикл. Вариант поведения хромосом, при котором разорванная хроматида соединяется с сестринской хроматидой, образуя гибрид. При расхождении центромер в ходе митоза, хромосома снова рвется (причем необязательно по месту соединения хромосом), тем самым возобновляя цикл.

Райзинг (W, число сверхвитков). Число самопересечений оси дуплекса в пространстве.

Рамка считывания. Один из трех возможных вариантов прочтения нуклеотидной последовательности. В каждой рамке считывания последовательность представлена как непрерывная серия триплетов. Любая последовательность заключает в себе три возможных рамки считывания. Выбор рамки считывания определяется выбором точки начала отсчета: если первая рамка начинается с позиции 1, то вторая рамка начинается с позиции 2, третья — с позиции 3 (позиция 4 представляет собой начало второго триплета первой рамки, и т.д.).

Ранние гены транскрибируются до репликации фаговой ДНК. Они кодируют регуляторы и другие белки, необходимые для последующих стадий инфекции.

Ранняя инфекция. Часть литического цикла бактериофага от попадания в клетку до начала репликации фаговой ДНК. За это время фаг синтезирует ферменты, необходимые для репликации его ДНК.

Рассеянные повторы. Первоначально были определены как короткие последовательности, в изобилии разбросанные по геному. В настоящее время установлено, что эти последовательности состоят из мобильных элементов.

р-белок, рибосомальный белок. Какой-либо из белков, входящих в состав рибосомы.

Ревертант. Мутантная клетка (организм), вернувшаяся к нормальному фенотипу.

Регион невосприимчивости (область иммунности). Сегмент генома бактериофага, позволяющий профагу ингибировать инфекцию своей бактерии еще одним фагом того же типа. Регион содержит ген, кодирующий белок-репрессор, а также сайты, с которыми связывается этот репрессор.

Регион переноса. Участок на F-плазмиде, необходимый для бактериальной коньюгации.

Редактирование РНК. Изменение кодирующей последовательности на уровне РНК после транскрипции.

Редкие мРНК. Совокупность большого числа разновидностей мРНК, каждая из которых представлена всего несколькими копиями на клетку. Именно эти молекулы задают высокую степень сложности пула последовательностей мРНК.

Резольваза. Фермент, который участвует в сайт-специфичной рекомбинации между двумя копиями транспозона, прошедшего дупликацию.

Рекомбинант «с заплаткой». Типичный результат одного из вариантов разделения структуры Холлидея — за счет внесения разрезов в перешедшие цепи. Одна из цепей дуплекса содержит «заплатку» — сегмент, пришедший из гомологичной хромосомы. Во всем остальном молекула та же, что до рекомбинации.

Рекомбинантное соединение. Место встречи двух рекомбинирующих дуплексов (с краю от гетеродуплексного участка).

Рекомбинационная репарация. Способ заполнения бреши в одной из цепей дуплекса с помощью привлечения гомологичной одиночной цепи из другого дуплекса.

Рекомбинационные узелки. Плотные объекты, наблюдаемые в синаптонемном комплексе; могут представлять собой белковые комплексы, участвующие в кроссинговере.

Релаксаза. Фермент, который разрезает одну из цепей ДНК и связывает образовавшийся 5′-конец.

Релизинг-фактор (RF) необходим для терминации синтеза белка. Он вызывает отделение законченной полипептидной цепочки и рибосомы от мРНК. Разновидности RF обозначаются цифрами. Эукариотические RF называются eRF и также обозначаются цифрами.

Ремоделирование хроматина. Термин обозначает энергетически зависимое смещение или реорганизацию нуклеосом в связи с активацией генов к транскрипции.

Ренатурация. Ассоциация денатурированных комплементарных одиночных цепей ДНК, снова образующих двойную спираль.

Репарация некомплементарностей. Исправление ошибок репликации в только что синтезированных участках дочерних дуплексов. Исправления вносятся преимущественно в дочернюю цепь (система различает дочернюю и матричную цепи: иногда по статусу метилирования).

Репарация. Восстановление интактной структуры ДНК после повреждения. Может выполняться как репаративный синтез (поврежденный участок цепи вырезается и заменяется новым материалом, синтезируемым на месте). В репарации могут также использоваться механизмы рекомбинации (двуцепочечный участок, содержащий повреждение, замещается интактным гомологичным участком с другой хромосомы).

Репликативная вилка. Точка разделения цепей матричного (родительского) дуплекса, где идет репликация. В репликативной вилке находится комплекс белков репликации, в том числе ДНК-полимераза.

Репликативная транспозиция. Перемещение транспозона согласно репликативному механизму, когда он сначала реплицируется, а затем одна из копий переносится в новый сайт.

Репликация двуцепочечной ДНК представляет собой синтез двух новых цепей, комплементарных цепям исходного (роди-

тельского) дуплекса. В результате родительский дуплекс превращается в пару идентичных дочерних дуплексов, каждый из которых содержит одну родительскую и одну новосинтезированную дочернюю цепь. Репликация называется полуконсервативной, поскольку половина (одна из цепей) каждого дочернего дуплекса представляет собой старый материал.

Репликон. Единица генома, в которой реплицируется ДНК. Каждый репликон содержит ориджин инициации репликации.

Реплисома. У бактерий: структура, содержащая множество белков, которая проходит сборку в репликативной вилке, чтобы вести синтез ДНК. Содержит ДНК-полимеразу и другие ферменты.

Репрессия. У бактерий репрессией называется подавление синтеза некоторых ферментов их собственными продуктами. В более широком смысле термин «репрессия» означает ингибирование транскрипции (или трансляции) посредством связывания белка-репрессора с конкретным сайтом в ДНК (или мРНК).

Репрессор. Белок, который ингибирует экспрессию гена. Некоторые репрессоры препятствуют транскрипции (посредством связывания с сайтом-оператором в ДНК), другие репрессоры препятствуют трансляции (связывая РНК).

Ретровирус. РНК-вирус, способный переводить свою геномную последовательность в форму ДНК при помощи обратной транскрипции.

Ретропозон. Транспозон, который перемещается в форме РНК. Элемент, встроенный в ДНК, подлежит транскрипции в РНК, а затем — обратной транскрипции в ДНК, которая встраивается в новый сайт генома. В отличие от ретровируса ретропозон не способен к инфекции.

Рецептор типа Toll (TLR). Рецептор в плазматической мембранефагоцитов и других клеток, которые участвуют в передаче сигнала при врожденном иммунном ответе. TLR похожи на рецепторы к IL-1.

Рибозим. РНК, обладающая каталитической активностью.

Рибонуклеазы. Ферменты, которые расщепляют РНК. Среди них есть специфичные к одно- или двуцепочечной РНК, эндо- и экзонуклеазы.

Рибонуклеопротеин. Комплекс РНК с белками.

Рибосома. Крупное образование, состоящее из РНК и белков, которое осуществляет синтез белка на матрице мРНК. Бактериальные рибосомы имеют коэффициент седиментации 70S, эукариотические рибосомы — 80S. Рибосома состоит из двух субъединиц.

Рибосомальная РНК (рибосомная РНК, рРНК). Главный компонент рибосомы. Обе субъединицы рибосомы содержат крупные рРНК, укомплектованные множеством белков.

РНК-лигаза. Фермент, который участвует в сплайсинге тРНК; образует фосфодиэфирную связь между двумя экзонами.

РНК-переключатель. Каталитическая РНК, активность которой зависит от взаимодействия с небольшим лигандом.

РНК-полимеразы. Ферменты, синтезирующие РНК на матрице ДНК (формальное название — ДНК-зависимые РНК-полимеразы).

Родительская, или матричная, цепь ДНК (родительский, или матричный, дуплекс ДНК). «Старые» цепи ДНК (в отличие от новосинтезированных). Термины используются при описании полуконсервативной репликации ДНК.

Ротационное позиционирование. Положение гистонового октамера относительно витков двойной спирали ДНК. От ротационного позиционирования зависит, какой стороной смотрит двойная спираль с поверхности нуклеосомы.

Ро-зависимые терминаторы. Последовательности, которые терминируют транскрипцию бактериальной РНК-полимеразой в присутствии ро-фактора.

Ро-фактор (ρ). Белок, который помогает РНК-полимеразе *E. coli* терминировать транскрипцию в некоторых терминаторах (такие терминаторы называются ро-зависимыми).

Сайленсер. Короткая последовательность ДНК, способная инактивировать экспрессию расположенного поблизости гена.

Сайленсинг РНК. Способность двухцепочечной РНК к систематическому подавлению экспрессии соответствующего гена в растительной клетке.

Сайленсинг (умолкание). Локальная репрессия работы гена, обычно связанная с изменением структуры хроматина.

Сайт Р. Сайт рибосомы, в котором находится пептидил-тРНК (молекула тРНК, несущая растущую полипептидную цепочку). В сайте Р пептидил-тРНК продолжает взаимодействовать с кодоном мРНК, узнанным ею в сайте А.

Сайт ветвления. Короткая последовательность, расположенная непосредственно перед окончанием интрона, по которой в ходе сплайсинга идет образование структуры типа лассо за счет соединения 5′-конца интрона и 2′-остатка аденозина.

Сайт прикрепления к матриксу (MAR). Участок геномной ДНК, который прикрепляется к ядерному матриксу. Также известен как «участок прикрепления к каркасу» (SAR).

Сайт связывания рибосомы. Последовательность в бактериальной мРНК, которая содержит инициирующий кодон, который связывается 30S субъединицей рибосомы на стадии инициации белкового синтеза.

Сайт сплайсинга. Последовательность в районе экзон-интронного стыка.

Сайт-специфичная рекомбинация. Рекомбинация между определенными последовательностями (например, при интеграции или вырезании фага или расхождении коинтегратов при транспозиции).

Сайты замен в гене. Сайты мутаций в кодирующей области, меняющих прочтение отдельных кодонов (что соответствует аминокислотным заменам в белке).

Самосборка. Принятие белком (или белковым комплексом) своей окончательной функциональной структуры без помощи дополнительных факторов (таких, как шапероны). Термин также применим к спонтанному образованию биологически-значимой структуры из молекул.

Сателлитная ДНК состоит из множества тандемных повторов короткой последовательности.

Сверхспирализация. Закрученность ковалентно замкнутй двуцепочечной ДНК таким образом, что она пересекает в пространстве собственную ось.

Свободный (неопределенный) сайт связывания. Любая произвольная последовательность ДНК, связываемая основным ферментом РНК-полимеразы, когда он не занят транскрипцией.

Сдвиги рамки. Мутации в кодирующей области, представляющие собой небольшие делеции или вставки, по размеру не кратные 3 пн. Они изменяют открытую рамку считывания, согласно которой триплеты транслируются в белок.

Сектор. Островок из клеток с необычным фенотипом, образованный потомством одной измененной клетки, в окружении обычных клеток той же ткани.

Семейство генов. Набор генов, дающих похожие или идентичные продукты (белки или РНК), в пределах одного генома. Семейства образуются за счет дупликаций предковых генов с последующим накоплением отличий в последовательностях копий. Чаще всего гены одного семейства похожи, но не идентичны.

Семейство последовательностей Alu. Набор разрозненных гомологичных последовательностей в геноме человека, длиной около 300 пн. Некоторые члены семейства несут по концам сайт узнавания для эндонуклеазы Alu (что и дало название семейству).

Септа. Структура, образующаяся посередине делящейся бактерии, в месте, где дочерние клетки будут отделены друг от друга. Термин «септа» также используется для обозначения клеточной стенки, образующейся между дочерними растительными клетками в конце митоза.

Септальное кольцо. Комплекс из нескольких белков, кодируемых генами *fts*, который образуется по середине длины клетки *E. coli*. Первоначальное называлось Z-кольцом, по названию белка FtsZ, который первым включается в его состав. Во время клеточного деления септальное кольцо дает начало септе.

Сигма-фактор (σ). Субъединица бактериальной РНК-полимеразы, необходимая для инициации; имеет огромное влияние на выбор промотора.

Сигнал-узнающая частица (SRP). Рибонуклеопротеиновый комплекс, узнающий сигнальную последовательность в самом начале процесса трансляции и обеспечивающий посадку рибосомы на канал мембраны ЭПР. Сигнал-узнающие частицы различных организмов могут иметь разную структуру, но в целом содержат гомологичные белки и РНК.

Сигнальная пептидаза. Фермент на мембране ЭПР, специфично отщепляющий сигнальную последовательность (сигнальный пептид) от белков непосредственно в процессе их транслокации через мембрану ЭПР. Аналогичные по активности ферменты есть у эубактерий, архебактерий и в органеллах эукариот, использующих мечение белков отщепляемыми последовательностями для целей их сортировки и транспорта. Сигнальная пептидаза — один из компонентов более крупного белкового комплекса.

Сигнальная последовательность. Короткий участок последовательности полипептида, обеспечивающий котрансляционную транслокацию несущего сигнальную последовательность белка внутрь ЭПР.

Сигнальный конец. Образуется при рекомбинации в генах иммуноглобулинов и Т-клеточных рецепторов. Сигнальные концы — это края фрагмента, содержащего сигнальные последовательности рекомбинации. От последующего соединения сигнальных концов возникает сигнальное соединение.

Синапсис. Ассоциация двух пар сестринских хроматид (представляющих собой гомологичные хромосомы), происходящая в начале мейоза. Образующуюся в результате структуру называют бивалентом.

Синаптонемный комплекс. Морфологическая структура находящихся в состоянии синапса хромосом.

Синонимичные кодоны. Кодоны, кодирующие одну и ту же аминокислоту. Синонимичные тРНК несут одну и ту же аминокислоту и взаимодействуют с одним и тем же кодоном.

Синтезируемая РНК. Рибонуклеотидная цепочка, находящаяся в состоянии синтеза; ее 3′-конец связан с ДНК, где идет элонгация РНК-полимеразой.

Синтения. Сходство хромосомных регионов различных видов, при котором гомологичные гены расположены в одинаковой ориентации.

Синтетическая летальность. Событие, при котором две мутации, каждая из которых сама по себе не летальна, в комбинации дают летальный фенотип.

Система жизненной зависимости. Механизм поддержания плазмиды в бактериальной популяции. убивающий бактерию в случае потери плазмиды.

Системы контроля. Системы ферментов, проверяющие нуклеиновые кислоты на наличие ошибок. Термин используют в нескольких значениях. В некоторых случаях им обозначают систему деградации мРНК, несущей нонсенс-мутации. Также термин используют для обозначения систем, реагирующих на повреждения структуры двойной спирали. Общим является использование данного обозначения систем, узнающих неправильные последовательности или структуры и запускающие механизм исправления ошибок.

Сквозное прочтение при транскрипции или трансляции происходит в случае, когда, соответственно, РНК-полимераза или рибосома игнорирует сигнал терминации. Возможные причины: мутация в матрице, неисправность вспомогательного фактора.

Скрытый сателлит. Разновидность сателлитной ДНК, не идентифицируемая как обособленный пик в градиенте плотности, т. е. входящая в состав ДНК «главной полосы».

Скрэпи (почесуха овец). Болезнь, вызываемая инфекционным агентом белковой природы (прионом).

Случайный дрейф. Флуктуация (без давления отбора) частот двух аллелей в популяции.

Совместная эволюция. Термин употребляется в ситуации, когда два гена подлежат эволюции как одно целое.

Согласованная эволюция. Термин употребляется для описания способности двух родственных генов к эволюции как будто в составе одного локуса.

Соединенная молекула. Пара дуплексов ДНК, образующих единую структуру за счет реципрокного обмена генетическим материалом.

Соединенный рекомбинант. ДНК, образующаяся в результате разводки структуры Холлидея путем разрезания не участвовавших в обмене цепей. В результате обе цепи ДНК до точки рекомбинации пришли из одной из гомологичных хромосом, а после точки рекомбинации — из другой.

Соматическая мутация. Мутация, возникающая в соматической клетке и затрагивающая лишь потомков данной клетки. Не передается по наследству.

Соматическая рекомбинация. Процесс соединения V-гена и C-гена в геноме лимфоцита для создания иммуноглобулина или Т-клеточного рецептора.

Сортировка белков. Распределение белков по направлениям транспорта в конкретные органеллы или между ними.

Спаривание оснований. Специфичное (комплементарное) взаимодействие аденина с тимином или гуанина с цитозином в двойной спирали ДНК (в двойной спирали РНК тимин замещен на урацил).

Спаривание хромосом. Расстановка хромосом парами (каждая со своим гомологом) в начале мейоза.

Спейсер. Последовательность внутри кластера генов, разделяющая повторяющиеся копии транскрипционной единицы.

Спираль называется правозакрученной, если витки огибают ось спирали по часовой стрелке.

Спираль узнавания. Одна из двух α-спиралей мотива «спираль-поворот-спираль», который устанавливает контакты с ДНК, специфичные к определенным сочетаниям оснований в нуклеотидной последовательности.

«Спираль-петля-спираль» (HLH). Мотив, обеспечивающий димеризацию; характерен для класса факторов транскрипции, известных как белки HLH. В белках bHLH рядом с этим мотивом находится положительно заряженный ДНК-связывающий мотив.

«Спираль-поворот-спираль». Мотив, образованный двумя α-спиралями, которые образуют сайт связывания ДНК. Одна из спиралей погружается в большую бороздку ДНК, вторая спираль ложится поперек.

Сплайсинг РНК. Этап процессинга транскриптов; можно представить как вырезание интронов и соединение экзонов в непрерывную мРНК.

Сплайсосома. Комплекс, формируемый рибонуклеопротеинами (в состав которых входят мяРНК) и наряду в другими белковыми факторами участвующий в сплайсинге.

Спонтанные мутации. Мутации, возникающие в отсутствие каких-либо специальных внешних воздействий, повышающих частоту возникновения мутаций. Является результатом ошибок репликации (или иных процессов, обеспечивающих воспроизводство ДНК).

Споруляция. Формирование спор бактериями (за счет морфологической перестройки) или дрожжами (в результате мейоза).

Средние гены (гены средней волны экспрессии). Гены бактериофага, регулируемые белками — продуктами ранних генов. Некоторые из белков, кодируемых средними генами, катализируют репликацию фаговой ДНК; другие средние гены регулируют экспрессию более поздних генов.

Стебель. Элемент структуры типа стебель-петля (шпилька) в РНК или ДНК.

Стероидные рецепторы. Транскрипционные факторы, активируемые в результате связывания со стероидным гормоном.

Стоп-кодон. Один из трех триплетов (UAG, UAA или UGA), вызывающих терминацию синтеза белка. Исторически они носят название бессмысленных (нонсенс) кодонов. Кодон UAA принято называть охра, а UAG — амбер (по названиям нонсенс-мутаций, при исследовании которых эти кодоны были впервые идентифицированы).

Структура типа лассо. Промежуточный продукт сплайсинга РНК, у которого есть «петля» (замкнутая связью 2'-5') и «хвост».

Структура Холлидея. Промежуточный продукт гомологичной рекомбинации, в котором два дуплекса ДНК соединены между собой генетическим материалом — двумя из четырех цепей, по одной из каждого дуплекса. Соединенная молекула разделяется внесением одноцепочечных разрезов (ников), которые восстанавливают существование двух раздельных дуплексов.

Структурный ген. Любой ген, кодирующий РНК или белок, не участвующие в регуляторных процессах.

Субвирусный патоген. Инфекционный агент, меньший по размеру, чем вирусы (например, вирусоид).

Супрессия. Генетический термин, означающий исчезновение действия мутации за счет некоторого генетического события, отличного от возврата к исходной последовательности ДНК.

Супрессор со сдвигом рамки. Вставка или делеция нуклеотида, восстанавливающая исходную рамку считывания в гене, утратившем активность в результате первичной вставки или делеции.

Супрессор. Повторная мутация, компенсирующая или предотвращающая эффект исходной мутации.

Сцепление. Тенденция двух генов наследоваться совместно, в виде устойчивого сочетания аллелей. Является результатом местонахождения этих генов на одной хромосоме; измеряется частотой рекомбинации между локусами, измеренной в процентах.

Талассемия. Болезнь крови, происходящая от отсутствия α- или β-глобина.

Т-ДНК. Сегмент Ti-плазмиды *Agrobacterium tumefaciens*, доставляемый в ядро растительной клетки в ходе инфекционного процесса. Несет гены, трансформирующие клетку растения.

Теломера. Концевая область хромосомы. Последовательность ДНК теломеры представляет собой простой повтор с выступающим одноцепочечным участком.

Теломераза. Рибонуклеопротеиновый фермент, синтезирующий повторяющиеся последовательности по концам хромосомы путем добавления нуклеотидов к 3'-концу ДНК в соответствии с последовательностью РНК-компонента рибонуклеопротеина.

Теломерный сайленсинг. Репрессия активности генов вблизи теломеры.

Тератома. Структура, в которой множество разных типов клеток (клетки кожи, зубов, костей и др.) начинают беспорядочный рост в случае трансплантации эмбриона на ранней стадии развития в одну из тканей взрослого животного.

Терминаза. Фермент, разрезает мультимеры вирусного генома и за счет использования энергии гидролиза АТФ обеспечивает перенос ДНК в пустой вирусный капсид начиная с разрезанного конца.

Терминальный белок. Белок, обеспечивающий репликацию линейного генома фага за счет затравления синтеза ДНК на самом конце молекулы. Ковалентно прикрепленный к 5'-концу геномной ДНК белок взаимодействует с ДНК-полимеразой и содержит остаток цитозина, выступающий в качестве затравки.

Терминатор (t). Последовательность ДНК, вызывающая терминацию транскрипции.

Терминация. Реакция окончания синтеза макромолекулы (репликации, транскрипции или трансляции) за счет прекращения добавления мономерных субъединиц и, как правило, вызывающая разборку синтезирующего аппарата.

Терминус. Участок ДНК, на котором заканчивается репликация.

Тест на комплементацию. Позволяет определить, в одном ли гене находятся две мутации. Тест выполняется путем скрещивания между носителями двух различных рецессивных мутаций с одинаковым фенотипом. Появление потомков с диким фенотипом говорит о принадлежности этих мутаций двум разным генам.

Тетрада. Четыре гаплоидных споры — продукт мейоза у дрожжей. (Первоначально этот термин обозначал структуру, характерную

для начала мейоза — бивалент, содержащий все четыре хроматиды, образованных дупликацией пары гомологичных хромосом.)

Т-клетки. Лимфоциты, происходящие из тимуса. Их подразделяют на несколько функциональных типов. Т-клетки несут TCR (Т-клеточный рецептор) и вовлечены в клеточно-опосредованный иммунный ответ.

Т-клеточный рецептор (TCR). Рецептор антигена на Т-лимфоците. Он клонально экспрессируется и связывается с комплексом белков МНС класса I или II и фрагментом антигена.

Толерантность. Отсутствие иммунного ответа на антиген вследствие клональной делеции.

Топоизомераза типа I. Фермент, изменяющий топологию ДНК путем внесения (и последующего сшивания) одноцепочечных разрывов.

Топоизомераза типа II. Фермент, изменяющий топологию ДНК путем внесения (и последующего сшивания) двуцепочечных разрывов.

Топоизомераза. Фермент, который изменяет число пересечений между двумя цепями в закрытой молекуле ДНК. Действует так: вносит в ДНК разрыв, пропускает ДНК через разрыв, «склеивает» концы.

Топологические изомеры. Молекулы ДНК, которые идентичны по всем признакам, кроме числа зацеплений.

Точечная мутация. Изменение последовательности ДНК, затрагивающее всего 1 пн.

Точка начала транскрипции. Позиция в ДНК, соответствующая первому нуклеотиду РНК.

Точная эксцизия. Выход полной последовательности транспозона из геномной ДНК (вместе с одной из двух копий последовательности-мишени). Это событие может восстановить функцию гена (промотора, регуляторной области и т.п.), в котором находилась вставка.

тпн (тысяча пар нуклеотидов, «килобаза»). Единица измерения длины ДНК или РНК.

Транзиция. Мутация, при которой пиримидин заменяется пиримидин или пурин на пурин.

Трансверсия. Мутация, при которой пурин заменяется на пиримидин (или наоборот).

Трансген. Введенный в клетку или целый организм чужеродный ген.

***Транс*-действующий продукт** (продукт, действующий в *транс*-манере) может работать с любой копией ДНК-мишени (значит, этот продукт представляет собой свободно диффундирующий белок или РНК).

Трансдуцирующий вирус. Вирус, захватывающий и переносящий часть генома клетки-хозяина вместо участка собственного генома. Наиболее известный пример — ретровирусы эукариот и ДНК-фаги *E. coli*.

***Транс*-конфигурация** двух сайтов означает их местонахождение на двух разных молекулах ДНК (например, хромосомах).

Транскрипт. РНК, продукт копирования одной из цепей ДНК. Транскрипт гена может нуждаться в процессинге для превращения в зрелую мРНК.

Транскриптом. Полный набор РНК, находящихся в данный момент в клетке, ткани или организме. Разнообразие транскриптома определяется в основном молекулами мРНК, однако также включает и некодирующие РНК.

Транскрипционная единица. Последовательность ДНК между сайтами инициации и терминации транскрипции. Может включать более одного гена.

Транскрипция. Синтез РНК на матрице ДНК.

Транслокация. Элемент импорта или экспорта белков или РНК через ядерную пору.

Транслокация белков. Перемещение белков через мембрану (мембрану органелл эукариот или плазматическую мембрану бактерий). Мембрана, через которую идет транслокация белков, обладает специально предназначенным для этого каналом.

Транслокация рибосомы. Перемещение рибосомы на один кодон по матрице мРНК после добавления очередной аминокислоты к синтезируемому полипептиду.

Транслокон. Трансмембранный белок, образующий канал, через который полипептид проходит сквозь мембрану.

Трансляционное позиционирование. Последовательное расположение гистоновых октамеров на соседних участках двойной спирали так, что они задают расположение определенных последовательностей ДНК в межнуклеосомных (линкерных) регионах.

Трансляция. Синтез белка по матрице мРНК.

Трансмембранный белок. Белок, пронизывающий мембрану насквозь. Гидрофобный регион (обычно представляющий собой последовательность из 20–25 гидрофобных или незаряженных аминокислот) или регионы белка находятся в толще мембраны. Гидрофильные области экспонированы по обеим сторонам мембраны.

Трансмембранный регион (трансмембранный домен). Участок белка, пронизывающий мембрану. Как правило, гидрофобен, и состоит из примерно 20 аминокислот, формирующих α-спираль.

Транспозаза. Энзиматическая активность, участвующая во вставке транспозона в новое место в геноме.

Транспозиция. Перемещение транспозона на новое место в геноме.

Транспозон. Последовательность ДНК, способная вставлять себя (или свою копию) в новое место в геноме при отсутствии какого-либо сходства транспозона с последовательностью целевого локуса.

Транспортная РНК (тРНК). Посредник при синтезе белков, интерпретирующий генетический код. тРНК может нести ковалентно присоединенную аминокислоту и имеет в своем составе антикодон, комплементарный кодирующему данную аминокислоту триплету на мРНК.

Трансфекция. Приобретение новых генетических маркеров эукариотическими клетками за счет поглощения ими внешней ДНК.

Трансформация бактерий. Приобретение нового генетического материала за счет поглощения внешней ДНК.

Трансформация эукариотических клеток. Переход клеток к неограниченному росту в культуре, схожий или идентичный опухолевому состоянию.

Трансформирующий элемент. ДНК, захватываемая бактерией, экспрессия которой меняет свойства клетки.

Трансэтерификация. Реакция скоординированного разрыва и воссоединения ковалентных связей таким образом, что дополнительной энергии для прохождения реакции не требуется.

тРНК$_f^{Met}$. Специальная (инициаторная) транспортная РНК в клетках бактерий, участвующая в доставке первой аминокислоты (формил-метионина) на рибосому и взаимодействующая со старт-кодоном (как правило, AUG, но иногда GUG или CUG).

тРНК$_i^{Met}$. Специальная (инициаторная) транспортная РНК в клетках эукариот, участвующая в доставке первой аминокислоты (метионина) на рибосому и взаимодействующая со старт-кодоном.

тРНК$_m^{Met}$. Транспортная РНК, присоединяющая метионин по внутренним кодонам AUG.

Тройственный комплекс. Комплекс, формирующийся при инициации транскрипции и состоящий из РНК-полимеразы, ДНК и динуклеотида, представляющего собой первые два нуклеотида синтезируемого транскрипта.

Трэйлер. Нетранслируемая 3′-концевая область мРНК, следующая за терминирующим кодоном.

Тяжелая цепь иммуноглобулина. Один из двух типов субъединиц тетрамера, которым является антитело. Каждое антитело содержит две тяжелых цепи. N-конец тяжелой цепи образует часть сайта узнавания антитела, С-конец определяет его изотип (принадлежность антитела к одному из подклассов).

Тяжелые цепи и легкие цепи дуплекса ДНК. Термин означает различие молекулярной массы комплементарных цепей ДНК, связанное с асимметричным составом азотистых оснований, так что одна цепь богата основаниями Т и G, другая богата основаниями С и А. Так бывает у некоторых сателлитных и митохондриальных ДНК.

Узел в ДНК. Запутанный участок, который не может быть разрешен в прямую молекулу без внесения разрезов и перестроек в ДНК.

Уникальная ДНК имеет кинетику реассоциации, отличную от ДНК, состоящей из повторов.

Усиливающая мутация в промоторе повышает интенсивность транскрипции.

Участки из вариабельного числа тандемных повторов (участки VNTR) в геноме содержат очень короткие повторяющиеся последовательности, включая микросателлиты и минисателлиты.

Участок контроля локуса (LCR). Участок, необходимый для экспрессии нескольких генов одного домена хромосомы.

Участок, богатый лейцинами. Мотив во внеклеточном домене некоторых поверхностных рецепторов в клетках растений и животных.

Фактор сборки. Белок, необходимый для формирования макромолекулярной структуры, но не входящий в состав данной структуры.

Фактор транскрипции (транскрипционный фактор). Фактор, который необходим РНК-полимеразе, чтобы начать транскрипцию с конкретного промотора; сам фактор не является частью фермента.

Факторы инициации (IF у прокариот, eIF у эукариот). Белки, ассоциированные с малой субъединицей рибосомы на стадии инициации синтеза белка.

Факторы элонгации (EF у прокариот, eEF у эукариот). Белки, которые циклически ассоциируют с рибосомами при добавлении каждой последующей аминокислоты к растущему полипептиду.

Факультативный гетерохроматин. Инертный хроматин, содержащий последовательности, способные существовать в активной форме — например, одна из Х-хромосом у самок млекопитающих.

Фиксация через кроссинговер. Возможное последствие неравного кроссинговера: мутация, возникшая в одном из генов тандемного кластера, распространяется на весь кластер (либо элиминируется).

Фиксация. Процесс замещения старого аллеля новым аллелем (в качестве аллеля, преобладающего в популяции).

Фоновый уровень мутаций. Скорость накопления изменений в последовательности генома данного организма. Является отражением баланса между возникновением спонтанных мутаций и их исправлением репарационными системами. Специфична для каждого вида организмов.

Фосфореле. Путь передачи фосфатной группы между сигнальными белками.

Фотореактивация. Путь расщепления циклобутан-пиримидиновых димеров, образовавшихся под влиянием ультрафиолета, при участии фермента, активируемого белым светом.

Фрагменты Оказаки. Короткие отрезки ДНК (по 1000-2000 нуклеотидов каждый) — промежуточный продукт полунепрерывной репликации отстающей цепи. В дальнейшем соединяются ковалентными связями в непрерывную цепь.

Футпринтинг. Метод идентификации сайтов связывания белков в ДНК, основанный на защищенности межнуклеотидных связей от атаки нуклеаз.

Хеликаза. Фермент, который разделяет цепи дуплекса нуклеиновой кислоты, используя энергию АТФ.

Хиазма. Участок видимого обмена генетическим материалом у двух гомологичных хромосом во время мейоза.

Химическая коррекция синтеза полимера. Механизм исправления ошибок синтеза, основанный на немедленном запуске обратной реакции в случае добавления к растущей цепи неправильного мономера.

Холофермент РНК-полимеразы. Форма РНК-полимеразы, компетентная к инициации транскрипции; состоит из четырех субъединиц основного фермента ($\alpha_2\beta\beta'$) и фактора σ.

Хоминг интронов. Способность некоторых интронов к внедрению в ДНК-мишень. Реакция специфична к уникальной последовательности-мишени.

Хроматиды. Копии одной хромосомы, продукт репликации, предназначенные для распределения по дочерним клеткам в последующем делении.

Хроматин. Термин описывает состояние ядерной ДНК и ассоциированных с ней белков во время интерфазы (между митозами) эукариотического клеточного цикла.

Хромомеры. Интенсивно окрашенные гранулы, которые видны на препаратах хромосом в определенных состояниях, особенно на ранних стадиях мейоза, когда хромосома может выглядеть как будто состоящей из серий таких хромомеров.

Хромомосомы типа ламповых щеток. Чрезвычайно распакованные мейотические биваленты, которые можно встретить в ооцитах некоторых амфибий.

Хромосома со множеством вилок (у бактерий) содержит более одного набора репликативных вилок в связи с тем, что вторая инициация произошла до завершения первого раунда репликации.

Хромосома. Дискретная единица генома, содержащая множество генов. Каждая хромосома состоит из очень длинной двухцепочечной молекулы ДНК, с которой ассоциирована примерно равная масса белков. Хромосому можно наблюдать как морфологическое целое лишь во время клеточного деления.

Хромоцентр. Представляет собой агрегат гетерохроматина из разных хромосом.

Центр организации микротрубочек. Область в цитозоле, в которой берут начало микротрубочки. В клетках животных главным организующим центром микротрубочек является центросома.

Центр сборки вируса табачной мозаики (ВТМ) представляет собой шпильку РНК, где инициируется сборка частицы (упаковка РНК субъединицами белка оболочки).

Центральная догма молекулярной биологии. Постулирует материальную природу сохранения, приумножения, передачи и реализации генетической информации. Последовательность мономеров в молекуле нуклеиновой кислоты подлежит воспроизведению за счет репликации, транскрипции или обратной транскрипции. Однако трансляция этой последовательности в аминокислотную последовательность полипептида (во время синтеза белков) является однонаправленной, потому что нуклеотидные последовательности нельзя выяснить, исходя из полипептидных последовательностей.

Центральная последовательность. Сегмент ДНК, присутствующий как в геноме фага лямбда, так и в геноме бактерии — в сайтах взаимодействия двух геномов. По этому сегменту проходит рекомбинация, приводящая к интеграции фага в бактериальную хромосому.

Центральный элемент. Структура, находящаяся посередине синаптонемного комплекса, вдоль которого происходит сопоставление латеральных элементов гомологичных хромосом. Образован белками Zip.

Центромера. Стянутый участок хромосомы, где находится место присоединения хромосомы к митотическому или мейотическому веретену (в этом месте хромосома образует специальную структуру, которая называется кинетохор). Центромера содержит уникальные последовательности ДНК и белки, отсутствующие в остальной части хромосомы.

Цинковый палец. ДНК-связывающий мотив, присущий факторам транскрипции соответствующего класса.

***цис*-конфигурация** означает, что два сайта принадлежат одной молекуле ДНК.

***цис*-действующий сайт** влияет на активность последовательностей ДНК (или РНК) в пределах одной молекулы. Обычно это указывает на то, что сайт не кодирует белок.

***цис*-доминантный сайт, или мутация,** определяет свойства лишь той молекулы ДНК, в которой сам находится. *цис*-доминантность указывает на отсутствие свободно диффундирующего белка, кодируемого в данном сайте. (Редкое исключение представляют собой *цис*-доминантные белки с эксклюзивной специфичностью к ДНК или РНК, на матрице которой они были синтезированы.)

Цистрон. Генетическая единица, определяемая по тестам на комплементацию; эквивалентна гену.

Цитоплазматический домен. Часть трансмембранного белка, погруженная в цитозоль.

Цитоплазматическое наследование. Обусловлено генами, находящимися в митохондриях и хлоропластах.

Цитотип. Состояние цитоплазмы, влияющее на активность Р-элемента у *D. melanogaster*. Сущность цитотипа заключается в наличии или отсутствии репрессора транспозиции, который достается яйцу от матери.

Цитотоксическая Т-клетка (Т-киллер). Т-лимфоцит (как правило, CD8+), который может быть стимулирован к уничтожению клеток, содержащих внутриклеточные патогены (например, вирусы).

Частицы-монстры бактериофага образуются из-за дефекта сборки капсида: он получается гораздо длиннее, чем нужно.

Число витков двойной спирали. Значение, соответствующее числу пар нуклеотидов на некотором участке ДНК, поделенному на число пар на один оборот спирали.

Число зацеплений. Число пересечений между двумя цепями ДНК в закрытом дуплексе.

Шапероны. Класс белков, связывающих не совсем зрелые (или не прошедшие сборку) белки для того, чтобы помочь их созреванию или чтобы помешать нежелательной агрегации незрелых белков.

Эволюционные часы определяются как скорость, с которой мутации накапливаются в отдельно взятом гене.

Эгоистичная ДНК. Геномные последовательности, не вносящие вклада в генотип; их единственная функция — поддерживать себя внутри генома.

Экзон. Сегмент прерывистого гена, представленный в зрелой мРНК.

Экзонуклеазы. Ферменты, которые отщепляют по одному нуклеотиду от конца полинуклеотидной цепочки. Эндонуклеаза может обладать специфичностью к 5′- или 3′-концу ДНК или РНК.

Экстеин. Сегмент аминокислотной последовательности, который сохраняется в зрелом белке, прошедшем белковый сплайсинг.

Эксцизионная репарация. Способ восстановления поврежденной ДНК, компетенция специальной системы со своим набором активностей. Принцип действия сводится к прямому уничтожению ущербного сегмента (цепи, содержащей повреждение) и его повторным синтезом этого сегмента на матрице интактной цепи.

Эксцизия (фага, эписомы или другой последовательности). Высвобождение из хромосомы хозяина в качестве автономной молекулы ДНК.

Эксцизия (этап эксцизионной репарации) представляет собой уничтожение сегмента поврежденной цепи ДНК под действием 5′–3′-экзонуклеазы.

Элемент ответа на глюкокортикоиды (GRE). Последовательность в промоторе или энхансере. Ее узнает рецептор глюкокортикоидов, активированный за счет взаимодействия со своим стероидным лигандом.

Элемент ответа на сыворотку (SRE). Последовательность в промоторе или энхансере, связываемая факторами транскрипции, индуцированными под воздействием сыворотки. Связыва-

ние SRE белком-активатором приводит к транскрипции генов, стимулирующих рост клеток.

Элемент ответа на тепловой шок (HSE). Последовательность в промоторе или энхансере, которая используется для активации гена регулятором, индуцированным тепловым шоком.

Элемент ответа. Последовательность в эукариотическом промоторе или энхансере, узнаваемая конкретным фактором транскрипции.

Элонгация. Стадия реакции макромолекулярного синтеза (репликации, транскрипции или трансляции), на которой нуклеотидная или полипептидная цепочка удлиняется за счет добавления новых субъединиц.

Эндонуклеазы рестрикции. Ферсменты, которые узнают в ДНК конкретные сайты-мишени (короткие последовательности) и разрезают дуплекс в этих сайтах (иногда — за их пределами, в зависимости от типа фермента).

Эндонуклеазы. Ферменты, которые расщепляют межнуклеотидные связи на расстоянии от концов полимера. Эндонуклеаза может обладать специфичностью к РНК, одно- или двухцепочечной ДНК.

Эндотоксин. Токсин, присутствующий на поверхности грамотрицательных бактерий (в отличие от экзотоксинов, секретируемых во внешнюю среду). Примером эндотоксина служит LPS.

Энхансер. *цис*-действующая последовательность, повышающая интенсивность использования эукариотического промотора. Энхансер может работать в любой ориентации, в любом местоположении (выше- или нижележащем) по отношению к промотору.

Энхансосома. Комплекс факторов транскрипции, проходящий кооперативную сборку на энхансере.

Эпигенетические изменения затрагивают фенотип без изменений в генотипе. Они представляют собой изменения в свойствах клетки; эти изменения наследуются, хотя не отображают никаких изменений генетической информации.

Эписома. Плазмида, способная к интеграции в бактериальную ДНК.

Эпитоп (антигенная детерминанта). Участок антигена, узнаваемый антигеновыми рецепторами на лимфоцитах.

Эпитоп. Часть антигена, распознаваемая рецептором на поверхности лимфоцитов. Другое название — антигенная детерминанта.

Эухроматин. Сложнейшая диффузная структура в интерфазном ядре; в составе эухроматина находится бóльшая часть геномной ДНК. Эухроматин менее компактен, чем гетерохроматин; он содержит большинство активных или потенциально активных генов, представленных в геноме одной копией.

Ядро промотора. Минимальная последовательность, на которой РНК-полимераза может инициировать транскрипцию (обычно с гораздо меньшей эффективностью, чем промотор, содержащий дополнительные элементы). Для РНК-полимеразы II ядро промотора представляет собой минимум, необходимый для сборки основного аппарата транскрипции, и включает в себя два элемента (InR и TATA-бокс). Характерная протяженность ядра промотора составляет 40 пн.

Ядрышко. Особая область эукариотического ядра, где продуцируются рибосомы.

Ядрышковое доминирование. Преобладание транскрипции генов рРНК, унаследованных от конкретного родителя, характерное для некоторых типов скрещивания.

Ядрышковый организатор. Область хромосомы, содержащая гены, кодирующие рРНК.

Якорь (часто упоминаемый как «трансмембранный якорь»). Участок трансмембранного белка, находящийся в мембране.

Предметный указатель

Н

О

П

Р

С

Т

Оглавление